

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:48:38 ; Search time 23.4674 Seconds
(without alignments)
791.964 Million cell updates/sec

Title: SEQ2_1-171_975-1163
Perfect score: 1823
Sequence: 1 MEDIDQSSSLVSSSTDSPPRP.....VKDAMAKIQAKIPGLKRKAD 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	904	49.6	199	2	US-08-700-607-1	Sequence 1, Appli
2	826.3	45.3	776	2	US-08-700-607-5	Sequence 5, Appli
3	790	43.3	356	2	US-08-700-607-6	Sequence 6, Appli
4	683	37.5	208	2	US-08-700-607-7	Sequence 7, Appli
5	669	36.7	267	2	US-08-700-607-8	Sequence 8, Appli
6	548.1	30.1	241	2	US-08-700-607-3	Sequence 3, Appli
7	539.9	29.6	168	4	US-09-149-476-563	Sequence 563, App
8	286	15.7	92	4	US-09-149-476-411	Sequence 411, App
9	262.5	14.4	75	4	US-09-621-976-4600	Sequence 4600, Ap
10	262.5	14.4	75	4	US-09-621-976-4601	Sequence 4601, Ap
11	253.2	13.9	1255	2	US-09-080-897-4	Sequence 4, Appli

12	253.2	13.9	1255	3	US-08-899-595-1	Sequence 1, Appli
13	253.2	13.9	1255	3	US-09-323-735-4	Sequence 4, Appli
14	252.5	13.9	1248	2	US-09-080-897-2	Sequence 2, Appli
15	252.5	13.9	1248	3	US-09-323-735-2	Sequence 2, Appli
16	248.8	13.6	1315	3	US-08-899-595-3	Sequence 3, Appli
17	236.6	13.0	933	3	US-08-764-870-14	Sequence 14, Appl
18	236.6	13.0	933	3	US-08-980-115-14	Sequence 14, Appl
19	236.2	13.0	1326	4	US-09-688-188B-15	Sequence 15, Appl
20	236.2	13.0	1326	4	US-09-291-417D-15	Sequence 15, Appl
21	235.7	12.9	8991	4	US-08-714-741-32	Sequence 32, Appl
22	230.6	12.6	1560	4	US-09-264-512B-2	Sequence 2, Appli
23	228.8	12.6	1070	3	US-08-922-635-22	Sequence 22, Appl
24	228.8	12.6	1504	4	US-09-364-206-2	Sequence 2, Appli
25	226.9	12.4	1187	1	US-08-320-559-28	Sequence 28, Appl
26	226.9	12.4	1187	3	US-08-545-860D-28	Sequence 28, Appl
27	226.9	12.4	1187	5	PCT-US94-04496-28	Sequence 28, Appl
28	226.9	12.4	1210	1	US-08-320-559-26	Sequence 26, Appl
29	226.9	12.4	1210	3	US-08-545-860D-26	Sequence 26, Appl
30	226.9	12.4	1210	5	PCT-US94-04496-26	Sequence 26, Appl
31	226.5	12.4	550	4	US-09-616-289-47	Sequence 47, Appl
32	225.7	12.4	1719	2	US-08-459-568-4	Sequence 4, Appli
33	225.7	12.4	1719	2	US-08-399-411-4	Sequence 4, Appli
34	225.7	12.4	1719	3	US-08-516-859A-4	Sequence 4, Appli
35	225.7	12.4	1719	4	US-09-586-472-4	Sequence 4, Appli
36	225.7	12.4	1719	4	US-09-528-706-4	Sequence 4, Appli
37	223.3	12.2	666	4	US-09-050-739-70	Sequence 70, Appl
38	221.8	12.2	826	4	US-09-894-998A-47	Sequence 47, Appl
39	221.1	12.1	830	4	US-09-562-737-31	Sequence 31, Appl
40	220.9	12.1	538	4	US-09-616-289-43	Sequence 43, Appl
41	220.9	12.1	830	4	US-09-562-737-35	Sequence 35, Appl
42	220.9	12.1	1912	4	US-08-913-832A-2	Sequence 2, Appli
43	220.9	12.1	1912	4	US-09-249-181A-2	Sequence 2, Appli
44	220.9	12.1	1912	4	US-09-158-707-2	Sequence 2, Appli
45	220.6	12.1	2482	4	US-09-252-991A-16967	Sequence 16967, A

ALIGNMENTS

RESULT 1

US-08-700-607-1

; Sequence 1, Application US/08700607

; Patent No. 5858708

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Au-Young, Janice

; APPLICANT: Goli, Surya K.

; APPLICANT: Hillman, Jennifer L.

; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,607
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 199 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-08-700-607-1

```

```

Query Match          49.6%; Score 904; DB 2; Length 199;
Best Local Similarity 97.3%; Pred. No. 1e-33;
Matches 183; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      173 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 232
          |||
Db      12 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 71

Qy      233 QKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVL 292
          |||
Db      72 QKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVL 131

Qy      293 MWVFTYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKI 352
          |||:|||||:|||||
Db      132 MWVFTYVGALENGLTLLILALISLFSVPVIYERHQVQIDHYLGLANKNVKDAMAKIQAKI 191

Qy      353 PGLKRKAD 360
          |||||:
Db      192 PGLKRKAE 199

```

RESULT 2

US-08-700-607-5

```

; Sequence 5, Application US/08700607
; Patent No. 5858708
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Au-Young, Janice
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.

```

```

; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,607
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 776 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 307307
US-08-700-607-5

```

```

Query Match          45.3%; Score 826.3; DB 2; Length 776;
Best Local Similarity 27.7%; Pred. No. 4.6e-29;
Matches 195; Conservative 52; Mismatches 91; Indels 367; Gaps 17;

```

```

QY      10 VSSSTDS-----PPR----- 19
      |||: |                ||:
Db      85 VSSAMDHTFSTTSKDGEGSCYTSLSIDICYPPQEDSTYFTGILQKENGHVTTISESPEELG 144

QY      20 -PPP-----AFKYQFVTEPED 34
      | |                |:| | :| |:
Db     145 TPGPSLPDVPGLIESRGLFSSDSGIEMTPAESTEVNKKILADPLDQMKAEAYKYIDITRPEE 204

QY      35 EEDEEEEEDEEED-----DED----- 50
      : |: | | | | | | | | | |
Db     205 VKHQEQHHPELEDKDLDFKNKDTDISIKPEGVREPDKPAPVEGKIIKDHLLEESTFAPYI 264

QY      51 -----LEELE----- 55
      | | |
Db     265 DDLSEEQRRAPIITTPVKITLTEIEPSVETTTQEKTPKQDICKPSPDTVPTVTVSEPE 324

QY      56 -----VLERKPAAGLS----- 66

```

```

                                : | | | |
Db      325 DDSPGSITPPSSGTEPSAAESQKGKSI SEDELITAIKEAKGLSYETAENPRPVGQLADRP 384

Qy      67 ---AAAVPPAAAAAPLLDFSSDSV-----PPA 89
          | : | | : | | | : |
Db      385 EVKARSGPPTIPSP-LDHEASSAESGDSEIELVSEDPMAAEDALPSGYVSFGHVGGPPPS 443

Qy      90 PRGPL-----PAAPPA----- 100
          | | |
Db      444 PASPSIQYSILREEREAE L DSELI IESCDASSASEESPKREQDSPMKPSALDAIREETG 503

Qy      101 --APERQPS-----WERSPAAPAPSLPPA-----AAVLPSKLPEDD----E 135
          | | | | : : | | | | : | | | | :
Db      504 VRAERAPSRRLAEPGSFLDYPSTEPQGPPELPPGDGALEPETPMLPRK-PEEDSSSNQ 562

Qy      136 PPARPPPPPPAGASPLAEP AAPPSTPAAPKRRGSGSSVVDLLYWRDIKKTGVVFGASLFL 195
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      563 SPAATKGPGLG-----PGAPPPLLFLNKQK-----AIDLLYWRDIKQTGIVFGSFLLL 611

Qy      196 LLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEEL 255
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      612 LFSLTQFSVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQ 671

Qy      256 VQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTTYVGALFNGLTLLILALIS 315
          : | | | : : | | | | | | | | | | | | | | | | | | | | | | | |
Db      672 IQKYTDCLQFYVNSTLKE L RRLFLVQDLVDLSLKFAVLMWLLTYVGALFNGLTLLLLMAVVS 731

Qy      316 LFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
          : | : | | | : | | | | | | : : | | | | | | | | | |
Db      732 MFTLPVVYVKHQAQIDQYLGLVRTHINAVVAKIQAKIPGAKRHA E 776

```

RESULT 3

US-08-700-607-6

; Sequence 6, Application US/08700607

; Patent No. 5858708

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Au-Young, Janice

; APPLICANT: Goli, Surya K.

; APPLICANT: Hillman, Jennifer L.

; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/700,607

```
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 307309
US-08-700-607-6
```

```
Query Match          43.3%; Score 790; DB 2; Length 356;
Best Local Similarity 44.1%; Pred. No. 4.6e-28;
Matches 177; Conservative 47; Mismatches 87; Indels 90; Gaps 15;
```

```
Qy      2 EDIDQSSLVS-SSTDSPPRPP--PAFKYQFVTEPEDEE-DEE-----EEEDEE 45
      || | || || | : : | : | : || | || :
Db      4 EDALPSGYVSFGHVGGPPPPSPASPSIQYSILREEREAEELDSIELIESCDASSASEESPKR 63

Qy     46 EDDLEDLEEVLERKPAAGLSAAAVPPAAAAPLLDFFSSDSV-----PPAPRGPL 94
      | | | | | | | | | | | | | | | | | : | | | | | | | | | | | |
Db     64 EQDS-----PPMKPSAL-----DAIREETGVRAEERAPSRRG-- 95

Qy     95 PAAPPAA----PERQPSWERSPAAPAPSLPPA-----AAVLPSKLPEDD----EPPAR 139
      | | : | : | | | | | | | | | | | | | | | | | | | | | | | |
Db     96 -LAEPGSFLDYPSTEPQ-----PGPELPPGDGALEPETPMLPRK-PEEDSSSNQSPAA 146

Qy    140 PPPPPPAGASPLAEPAPAPPSTPAAPKRRGSGSSVVDLLYWRDIKKTGVVFGASLFLLLSL 199
      | | | | | | | | | | | | | | | | | : | | | | | | | | | | | |
Db    147 TKGPGPLG-----PGAPPPLLFLNKQK-----AIDLLYWRDIKQTGIVFGSFLLLLFSL 195

Qy    200 TVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKY 259
      | | : | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db    196 TQFSVVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEIQKY 255

Qy    260 SNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTTYVGALFNGLTLLILALISLFSI 319
      : : : | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db    256 TDCLQFYVNSTLKLRLFLVQDLVDLSLKFAVLMWLLTYVGALFNGLTLLILMAVVSMTL 315

Qy    320 PVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
      | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    316 PVVYVKHQAQIDQYLGLVLRTHINAVVAKIQAKIPGAKRHAE 356
```

```
RESULT 4
US-08-700-607-7
; Sequence 7, Application US/08700607
; Patent No. 5858708
```

```

; GENERAL INFORMATION:
;   APPLICANT:  Bandman, Olga
;   APPLICANT:  Au-Young, Janice
;   APPLICANT:  Goli, Surya K.
;   APPLICANT:  Hillman, Jennifer L.
;   TITLE OF INVENTION:  TWO NOVEL HUMAN NSP-LIKE PROTEINS
;   NUMBER OF SEQUENCES:  9
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;     STREET:    3174 Porter Drive
;     CITY:      Palo Alto
;     STATE:     CA
;     COUNTRY:   U.S.
;     ZIP:       94304
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Diskette
;     COMPUTER:     IBM Compatible
;     OPERATING SYSTEM:  DOS
;     SOFTWARE:     FastSEQ Version 1.5
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/700,607
;     FILING DATE:       Filed Herewith
;   ATTORNEY/AGENT INFORMATION:
;     NAME:              Billings, Lucy J.
;     REGISTRATION NUMBER:  36,749
;     REFERENCE/DOCKET NUMBER:  PF-0114 US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:         415-855-0555
;     TELEFAX:           415-845-4166
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:          208 amino acids
;       TYPE:             amino acid
;       STRANDEDNESS:     single
;       TOPOLOGY:         linear
;     MOLECULE TYPE:      peptide
;     IMMEDIATE SOURCE:
;       LIBRARY:          GenBank
;       CLONE:            307311
US-08-700-607-7

```

```

Query Match          37.5%;  Score 683;  DB 2;  Length 208;
Best Local Similarity 67.4%;  Pred. No. 1.2e-23;
Matches 128;  Conservative 31;  Mismatches 31;  Indels 0;  Gaps 0;

```

```

Qy      171 SSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIO 230
      | :|||||||:|:|:|: | || ||| ||:| | |:| | ||||| | :|
Db      19  SQAIDLlyWRDIKQTGIVFGSFLLLSLTQFSVSVVAYLALAALSATISFRIYKSVLQ 78

Qy      231 AIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFA 290
      |:|:|:|:|:|:|:| | | : :|:| :|:|:| :|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      79  AVQKTDEGHPFKAYLELEITLSQEIQKYTDCLQFYVNSTLKELRRLFLVQDLVDLSLKFA 138

Qy      291 VLMWVFTYVGALFENGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQA 350
      |||:| | |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      139 VLMWLLTYVGALFENGLTLLMAVVMFTLPVVVYVKHQAQIDQYLGLVRTHINAVVAKIQA 198

```

Qy 351 KIPGLKRRKAD 360
|||| || |:
Db 199 KIPGAKRHAE 208

RESULT 5

US-08-700-607-8

; Sequence 8, Application US/08700607
; Patent No. 5858708
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Au-Young, Janice
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,607
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 281046
US-08-700-607-8

Query Match 36.7%; Score 669; DB 2; Length 267;
Best Local Similarity 66.7%; Pred. No. 8.1e-23;
Matches 124; Conservative 32; Mismatches 30; Indels 0; Gaps 0;

Qy 171 SSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQ 230
| :|||||||:|:|:| | || ||| ||:| | | :| | | | | | | | | | :|
Db 10 SQAIDLLYWRDIKQTGIVFGSFLLLSLTQFSVSVVAYLALAALSATISFRIYKSVLQ 69

US-08-700-607-3

Query Match 30.1%; Score 548.1; DB 2; Length 241;
Best Local Similarity 44.4%; Pred. No. 2e-17;
Matches 111; Conservative 40; Mismatches 60; Indels 39; Gaps 3;

```
Qy      109 ERSPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRG 168
      || | : | : :: | | | | | | | | | |
Db      3 ERXAATQSHSISSSSF-----GAEPSAPGGGGSPGACPA-----LGTKSCS 43

Qy      169 SGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 228
      | : | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      44 SSCAVHDLIXWRDVKKTFGVFGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRIYKSV 103

Qy      229 IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDILVDSLK 288
      | | : | | : | | | | : | | : : | | | | | | | | | | | | | | | | | | | | | |
Db      104 IQAVQKSEEGHPFKAYLDVDITLSSEAFHNYMNAAMVHINRALKLIIRLFLVEDILVDSLK 163

Qy      289 FAVLMWVFITYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKI 348
      | | | | : | | | | : | | | | | | | | : : | : | : | : | : | : | : | : | : | :
Db      164 LAVFMWLMITYVGAVFNGITLLILAELLIXSVPIVYXKY-----KV 203

Qy      349 QAKIPGLKRK 358
      : | | : : |
Db      204 PSKTPWNRQK 213
```

RESULT 7

US-09-149-476-563
; Sequence 563, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312

; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060

; EARLIER FILING DATE: 1997-10-02

Query Match 29.6%; Score 539.9; DB 4; Length 168;
Best Local Similarity 59.3%; Pred. No. 2.4e-17;
Matches 99; Conservative 36; Mismatches 31; Indels 1; Gaps 1;

```
Qy      195 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEE 254
          :|||| ||:| | :| :| ||||| ||||| ||||| ||||| :| :| |
Db      1 MLLSLAAFSVISVVSYLILALLSVTISFRIYKSVIQAVQKSEEGHPFKAYLDVDITLSSE 60

Qy      255 LVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGFALFNGLTLLILALI 314
          | |:| :| | :| :| ||||| ||||| || ||: ||||| ||||| :
Db      61 AFHNYMNAAMVHINRALKLIIRLFLVEDLVDSLKLAVFMWLMTYVGAVFNGITLLILAE 120

Qy      315 SLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGL-KRKAD 360
          :||:|:|:|:|:| |||||:| | : |||||:|:| :|:|:
Db      121 LIFSVPIVYEKYKTQIDHYVGIARDQTKSIVEKIQAKLPGIAKKKAE 167
```

RESULT 8

US-09-149-476-411

; Sequence 411, Application US/09149476

; Patent No. 6420526

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 186 Human Secreted proteins

; FILE REFERENCE: PZ002P1

; CURRENT APPLICATION NUMBER: US/09/149,476

; CURRENT FILING DATE: 1998-09-08

; EARLIER APPLICATION NUMBER: PCT/US98/04493

; EARLIER FILING DATE: 1998-03-06

; EARLIER APPLICATION NUMBER: 60/040,162

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,333

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/038,621

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,626

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,334

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,336

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,163

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/047,600

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,615

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,597

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,502

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,633

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,583

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23

Db 2 NAAMVHINRALKLIIRLFLVEDLVDSLKLAVFMWLMITYVGAVFNGITLLILAELLIFSVP 61

Qy 321 VIYERHQVQIDHYLGLANKSVKDAMAKIQAK 351
 ::||::: |||||:|:| | : || :|

Db 62 IVYEKYKTQIDHYVGIARDQTKSIVEKIPSK 92

RESULT 9

US-09-621-976-4600
 ; Sequence 4600, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 4600
 ; LENGTH: 75
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: 58
 ; OTHER INFORMATION: Xaa = His,Pro
 ; NAME/KEY: UNSURE
 ; LOCATION: 28
 ; OTHER INFORMATION: Xaa = Met,Val
 ; NAME/KEY: UNSURE
 ; LOCATION: 19
 ; OTHER INFORMATION: Xaa = Pro,Gln
 ; NAME/KEY: UNSURE
 ; LOCATION: 53
 ; OTHER INFORMATION: Xaa = Ser,Tyr
 US-09-621-976-4600

Query Match 14.4%; Score 262.5; DB 4; Length 75;
 Best Local Similarity 67.0%; Pred. No. 2e-05;
 Matches 59; Conservative 3; Mismatches 11; Indels 15; Gaps 3;

Qy 1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEEELEVLERK 60
 |||:|||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1 MEDLDQSPVSSS-DSPPRXQPAFKYQFXREPEDEE-----EDLEEELEVLERK 47

Qy 61 PAAGLSAAAV--PPAAAAPLLDFSSDSV 86
 ||||| || | ||| |||:| | :| |

Db 48 PAAGLXAAPVXTAPAAGAPLMDFGNDFV 75

RESULT 10

US-09-621-976-4601
 ; Sequence 4601, Application US/09621976
 ; Patent No. 6639063

```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4601
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 58
; OTHER INFORMATION: Xaa = His,Pro
; NAME/KEY: UNSURE
; LOCATION: 28
; OTHER INFORMATION: Xaa = Met,Val
; NAME/KEY: UNSURE
; LOCATION: 19
; OTHER INFORMATION: Xaa = Pro,Gln
; NAME/KEY: UNSURE
; LOCATION: 53
; OTHER INFORMATION: Xaa = Ser,Tyr
US-09-621-976-4601
```

```
Query Match          14.4%; Score 262.5; DB 4; Length 75;
Best Local Similarity 67.0%; Pred. No. 2e-05;
Matches 59; Conservative 3; Mismatches 11; Indels 15; Gaps 3;
```

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEDEDEDEDEDEDEDEDEDEEEDLEEVLERK 60
      |||:|||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 MEDLDQSPLVSSS-DSPPRXQPAFKYQFXREPEDEE-----EDLEEVLERK 47

Qy      61 PAAGLSAAAV--PPAAAAPLLDFSSDSV 86
      ||||| || | ||| |||:|| :| |
Db      48 PAAGLXAAPVXTAPAAGAPLMDFGNDFV 75
```

RESULT 11

US-09-080-897-4

```
; Sequence 4, Application US/09080897
; Patent No. 5985574
; GENERAL INFORMATION:
; APPLICANT: King, Mary-Claire
; APPLICANT: Lynch, Eric D.
; APPLICANT: Lee, Ming
; APPLICANT: Morrow, Jan E.
; APPLICANT: Welcsh, Piri L.
; APPLICANT: Leon, Pedro E.
; TITLE OF INVENTION: Modulators of Actin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
```

```

; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,897
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UW97-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-080-897-4

```

```

Query Match          13.9%; Score 253.2; DB 2; Length 1255;
Best Local Similarity 16.7%; Pred. No. 0.012;
Matches 142; Conservative 43; Mismatches 126; Indels 538; Gaps 36;

```

```

Qy      2 EDIDQSSILVSSSTDSPRPAPAFKYQFV-----TEPEDEED 37
      | : | | : | | | | | : : | | | : |
Db      427 ECVSQIVLHKNGTD-----PDFKCRHLQIDIERLVDQMIDKTKVEKSEAKATELEKKLD 480

Qy      38 EE-----EEDEEEDDEDLE-----ELEV----- 56
      | | | | : : | | | | | | | | | |
Db      481 SELTARHELQVEMKKMENDFEQKLQDLQGEKDALDSEKQQITAQKQDLEAEVSKLTGEVA 540

Qy      57 -----LE--RKPAAGLSA-----AAVPPAAAAPLLDFFSSDSV-----PPAPRG 92
      | | : | | | | | | | | | | | | | | | |
Db      541 KLSKELEDAKNEMASLSAVVVAPSVSSSAAVPP---APPLPGDSGTVIPPPPPPPPLPGG 597

Qy      93 PLPAAPPAAP-----ERQPSWERSPAAPAP-----SLPP----- 121
      : | : | | | | | | | | | | | | | | | |
Db      598 VPPSPPLPPGTCIPPPPLPGGACIPPPPLPGSAAIPPPPLPGVASIPPPPLPGAT 657

Qy      122 -----AAVLV-----SKLPEDDEP-----PARPPPPPPAGA 148
      | | : | : : | | | | | | | | | | | |
Db      658 AIPPPPLPGATAIPPPPLPGGTGIPPPPLPGSVGVPPPPPLPGGPGLPPIPPPPFPG 717

Qy      149 SP-----LAEPAPAP---STPAAP-----KRRGSGSSV---DL--- 176
      : | : | | | | | | | | | | | | | | | |
Db      718 APGIPPPPGMGVPPPPPGFGVPAAPVLPFGLTPKKVKYKPEVQLRRPNWSKFVAEDLSQ 777

```

Qy 177 -LYWRDIK----- 183
 :| :|
 Db 778 DCFWTKVKEDRFENNELFAKLTIAFSAQTKTSKAKKDQEGGEEKKSVQKKVKELKVLDS 837

 Qy 184 KT----- 185
 ||
 Db 838 KTAQNLSIFLGSRMPYQEIKNVILEVNEAVLTESMIQNLIKQMPPEQLKMLSELKEEY 897

 Qy 186 -----GVVFG-----ASLFLLLSLTVFS-----IVSVTA----- 209
 ||| | | | | | || | | | |
 Db 898 DDLAEESEQFGVVMGTVPRLRPRLNAILFKL----QFSEQVENIKPEIVSVTAACEELRKS 953

 Qy 210 -----YIALALL-----SVTISF----RIYKGVIAI-----QKSDEG 238
 : | | | | | | | | | : : :
 Db 954 ENFSSLLELTLLVGNMAGSRNAGAFGFNISFLCKLRDTKSADQKMTLLHFLAELCEND 1013

 Qy 239 HP-----FRAYLESEVAISEELVQK----- 258
 || | : : | : | | : | |
 Db 1014 HPEVLKFPDELAHVEKASRVSAENLQKSLDQMKKQIADVERDVQNFPAAATDEKDKFVEKM 1073

 Qy 259 -----YSNSALGHVN--STIKEL-----RRLFL---- 279
 | : : | | : | | | | : | |
 Db 1074 TSFVKDAQEQYNKLRMMHNSMETLYKELGDYFVFDPKKLSVEEFFMDLHNFRNMFLQAVK 1133

 Qy 280 -----VDDLVDSLKFAV 291
 : | | : : | :
 Db 1134 ENQKRRETEEKMRRAKLAKAEKERLEKQQKREQLIDMNAEGDETGVMDSLLEALQ--- 1190

 Qy 292 LMWVFTTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLG LANKSV-----KDA 344
 || | : | | : : | | | : ||
 Db 1191 -----SGAAFR-----RKRGRQVNRKAGCAVTSLLASELTKDDA 1225

 Qy 345 MAKIQAKIP 353
 || | : |
 Db 1226 MAPGPVKVP 1234

RESULT 12

US-08-899-595-1

; Sequence 1, Application US/08899595

; Patent No. 6111072

; GENERAL INFORMATION:

; APPLICANT: Narumiya, Shuh

; APPLICANT: Takahashi, No. 6111072uaki

; TITLE OF INVENTION: RHO TARGET PROTEIN HUMAN MDIA AND GENE

; TITLE OF INVENTION: ENCODING SAME

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

```

;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/899,595
;   FILING DATE:  24-JUL-1997
;   CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  JP 8-242701
;   FILING DATE:  26-AUG-1996
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  JP 9-90170
;   FILING DATE:  25-MAR-1997
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  Stephen A. Bent
;   REGISTRATION NUMBER:  29,768
;   REFERENCE/DOCKET NUMBER:  049441/0112
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (202)672-5300
;   TELEFAX:  (202)672-5399
;   TELEX:  904136
;   INFORMATION FOR SEQ ID NO:  1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  1255 amino acids
;   TYPE:  amino acid
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-08-899-595-1

```

```

Query Match          13.9%;  Score 253.2;  DB 3;  Length 1255;
Best Local Similarity  16.7%;  Pred. No. 0.012;
Matches 142;  Conservative 43;  Mismatches 126;  Indels 538;  Gaps 36;

```

```

Qy      2 EDIDQSSILVSSSTDSPRRPPPAFKYQFV-----TEPEDEED 37
      | : | | : ||      | || : :      || | : |
Db      427 ECVSQIVLHKNGTD-----PDFKCRHLQIDIERLVDQMIDKTKVEKSEAKATELEKKLD 480

Qy      38 EE-----EEEDEEEDDEDLE-----ELEV----- 56
      |      | | | : : || :      | ||
Db      481 SELTARHELQVEMKKMENDFEQKLQDLQGEKDALDSEKQQITAQKQDLEAEVSKLTGEVA 540

Qy      57 -----LE--RKPAAGLSA-----AAVPPAAAAPLLDFSSDSV-----PPAPRG 92
      || : | |||      |||| | | | | : |      || | |
Db      541 KLSKELEDAKNEMASLSAVVAVPSVSSSAAVPP---APPLPGDSGTVIPPPPPPPPLPGG 597

Qy      93 PLPAAPPAAP-----ERQPSWERSPAAPAP-----SLPP----- 121
      : | : || |      |      | | | |      | : ||
Db      598 VVPPSPPLPPGTCIPPPPPPLPGGACIPPPPLPGSAAIPPPPLPGVASIPPPPPPLPGAT 657

Qy      122 -----AAAVLP-----SKLPEDDEP-----PARPPPPPPAGA 148
      | | : |      : : |      |      | | ||||
Db      658 AIPPPPPPLPGATAIPPPPPPLPGGTGIPPPPPPLPGSVGVPPPPPLPGGPGLPPIPPPPPPFPG 717

Qy      149 SP-----LAEPAAPP---STPAAP-----KRRGSGSSVV--DL-- 176
      : |      : | ||      ||||      || : | | ||
Db      718 APGIPPPPPGMGVPPPPPPFGFGVPAAPVLPFGLTPKKVYKPEVQLRRPNWSKFVAEDLSQ 777

```

Qy 177 -LYWRDIK----- 183
 :| :|
 Db 778 DCFWTKVKEDRFENNELFAKLTlafsaQTKTSKAKKDQEGGEEKSVQKKVKELKVLDs 837

 Qy 184 KT----- 185
 ||
 Db 838 KTAQNLSIFLGsFRMPYQEIKNVILEVNEAVLTESMIQNLIKQMPePEQLKMLSELKEEY 897

 Qy 186 -----GVVFG-----ASLFLLLSLTVFS-----IVSVTA----- 209
 ||| | | | | | || |||||
 Db 898 DDLAESEQFGVVMGTVPRLRPRLNAILFKL---QFSEQVENIKPEIVSVTAACEELRKS 953

 Qy 210 -----YIALALL-----SVTISF-----RIYKGVIQAI-----QKSDEG 238
 : | || ||| | | | : : :
 Db 954 ENFSSLLELTLLVGNMAGSRNAGAFGFNISFLCKLRDTKSADQKMTLLHFLAELCEND 1013

 Qy 239 HP-----FRAYLESEVAISEELVQK----- 258
 || |: :| : | | : ||
 Db 1014 HPEVLKFPDELAHVEKASRVSAENLQKSLDQMKKQIADVERDVQNFPAAATDEKDKFVEKM 1073

 Qy 259 -----YSNSALGHVN--STIKEL-----RRLFL---- 279
 |: : | | : ||| | : ||
 Db 1074 TSFVKDAQEQYNKLRRMMHSNMETLYKELGDYFVFDPKKLSVEEFFMDLHNFRNMFLQAVK 1133

 Qy 280 -----VDDLVDSLKFAV 291
 : | | : : | :
 Db 1134 ENQKRRETEEKMRRAKLAKKEKAERLEKQKQKREQLIDMNAEGDETGVMDSLLEALQ--- 1190

 Qy 292 LMWVFYTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSV-----KDA 344
 || | : | | : : | | | : ||
 Db 1191 -----SGAAFR-----RKRGRQVNRKAGCAVTSLLASELTKDDA 1225

 Qy 345 MAKIQAKIP 353
 || | : |
 Db 1226 MAPGPVKVP 1234

RESULT 13

US-09-323-735-4

; Sequence 4, Application US/09323735

; Patent No. 6197932

; GENERAL INFORMATION:

; APPLICANT: King, Mary-Claire

; APPLICANT: Lynch, Eric D.

; APPLICANT: Lee, Ming

; APPLICANT: Morrow, Jan E.

; APPLICANT: Welcsh, Piri L.

; APPLICANT: Leon, Pedro E.

; TITLE OF INVENTION: Modulators of Actin

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 75 DENISE DRIVE

; CITY: HILLSBOROUGH

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94010

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/323,735
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/080,897
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UW97-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-323-735-4

```

```

Query Match          13.9%; Score 253.2; DB 3; Length 1255;
Best Local Similarity 16.7%; Pred. No. 0.012;
Matches 142; Conservative 43; Mismatches 126; Indels 538; Gaps 36;

```

```

QY      2 EDIDQSSLVSSSTDSPRPPPAFKYQFV-----TEPEDEED 37
      | : | | : ||      | || : :      || | : |
Db      427 ECVSQIVLHKNGTD-----PDFKCRHLQIDIERLVDQMIDKTKVEKSEAKATELEKKLD 480

QY      38 EE-----EEDEEEDDEDLE-----ELEV----- 56
      |      | | | : : || :      | ||
Db      481 SELTARHELQVEMKKMENDFEQKLQDLQGEKDALDSEKQQITAQKQDLEAEVSKLTGEVA 540

QY      57 -----LE--RKPAAGLSA-----AAVPPAAAAPLLDFSSDSV-----PPAPRG 92
      || : | |||      |||| | | | : |      || | |
Db      541 KLSKELEDAKNEMASLSAVVAPSVSSSAAVPP---APPLPGDSGTVIPPPPPPPPLPGG 597

QY      93 PLPAAPPAAP-----ERQPSWERSPAAPAP-----SLPP----- 121
      : | : || |      |      | | | |      | : ||
Db      598 VVPPSPPLPPGTCIPPPPPPLPGGACIPPPPPQLPGSAAIPPPPPPLPGVASIPPPPPPLPGAT 657

QY      122 -----AAAVLP-----SKLPEDDEP-----PARPPPPPPAGA 148
      | | : |      : : |      |      | | ||||
Db      658 AIPPPPPPLPGATAIPPPPPPLPGGTGIPPPPPPLPGSVGVPPPPPLPGGPGLPPIPPPPPPFPG 717

QY      149 SP-----LAEPAAPP---STPAAP-----KRRGSGSSVV--DL-- 176
      : |      : | ||      ||||      || : | | ||
Db      718 APGIPPPPPGMGVPPPPPPFGFGVPAAPVLPFGLTPKKVYKPEVQLRRPNWSKFVAEDLSQ 777

QY      177 -LYWRDIK----- 183

```



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,897
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UW97-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1248 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-080-897-2

```

```

Query Match          13.9%; Score 252.5; DB 2; Length 1248;
Best Local Similarity 15.3%; Pred. No. 0.012;
Matches 140; Conservative 46; Mismatches 102; Indels 625; Gaps 39;

```

```

Qy      1 MEDIDQ--SSLVSSSTDSPPRP----- 20
      |:| :: |::: || |
Db      378 MDDFNEVFQILLNTVKDSKAEPHFLSILQHLLLVVRNDYEARPQYYKLIEECISQIVLHKN 437

Qy      21 --PPAFKY-----QFVTEPEDEEDEEEE-----E 41
      | || | : : : | : | :
Db      438 GADPDFKCRHLQIEIEGLIDQMIDKTKVEKSEAKAAELEKKLDSELTARHELQVEMKKME 497

Qy      42 EDEEEDDEDLE-----ELEVLE-----RKPAAGLS 66
      | |: :||: | || : :| | ||
Db      498 SDFEQKLQDLQGEKDALHSEKQQIATEKQDLEAEVSQLTGEVAKLTKELEDACKEMASLS 557

Qy      67 AAA--VPPA--AAAPLLDFSSDSVPPAPRGPLPA-----APPAAP--ERQPSWERSPAA 114
      ||| |||: : || |||| ||| |||| | |
Db      558 AAATVPPSVPSRAP-----VPPAP--PLPGDSGTIIPPPAPGDSTTPPPPPPPPP 607

Qy      115 PAPSLPPAAAVLP-----SKLP----- 131
      | | || |: | | ||
Db      608 PPPPLPGGTALSPPPPLSGDATIPPPPLPEGVGIPSPSSLPGGTAIPPPPLPGSARIP 667

Qy      132 -----EDDEPPARPP-----PPPPAGASPL 151
      | || || | ||| | ||| :
Db      668 PPPPLPGSAGIPPPPPPLPGEAGMPPPPPLPGGPGIPPPPPFPGGPGIPPPPPG---M 724

Qy      152 AEPAAPP---STPAAP-----KRRGSGSSVV--DL---LYWRDIK--- 183
      | || |||| || : | :| || :| :|
Db      725 GMPPPPPFGEFGVPAAPVLPFGLTPKKLYKPEVQLRRPNWSKLVAEDLSQDCFWTKVKEDR 784

Qy      184 -----KT----- 185
      ||

```

Db 785 FENNELFAKLTLTFSAQTKTKKDQEGGEEKKSVQKKVKELKVLD SKTAQNLSIFLG SFR 844
 Qy 186 -----G VVFG 190
 ||| |
 Db 845 MPYQEIKNVILEVNEAVLTESMIQNLIKQMPPEPEQLKMLSELKDEYDDLAESEQFGVVMG 904
 Qy 191 -----ASLFLLLSLTVFS-----IVSVTAYIALA-----LLSVT 219
 | || | || ||||| | || :|
 Db 905 TVPRLRPRLNAILFKL----QFSEQVENIKPEIVSVTA----ACEELRKSESFSNLEIT 956
 Qy 220 I-----SFRI-----YKGVIQAIQK 234
 : | | | : :
 Db 957 LLVGNYMNAGSRNAGAFGFNISFLCKLRDTKSTDQKMTLLHFLAELCENDYPDVLKF--- 1013
 Qy 235 SDEGHPFRAYLESEVAISEELVQK----- 258
 || | :| :| | :||
 Db 1014 PDE----LAHVEKASRVSAENLQKNLDQMKKQISDVERDVQNFPAAATDEKDKFVEKMTSF 1069
 Qy 259 -----YSNSALGHVN--STIKEL-----RRLFL----- 279
 | : : | | : ||| | :||
 Db 1070 VKDAQEQYNKL RMMHSNMETLYKELGEYFLFDPKKLSVEEFFMDLHNFRNMFLQAVKENQ 1129
 Qy 280 -----VDDLVD SLKFAVLMW 294
 :| | : :| :
 Db 1130 KRRKTEEKMRRAKLAKAEKERLEKQOKREQLIDMNAEGDETGVMDSLLEALQ----- 1183
 Qy 295 VFTYVGALFN-----GLTLLILALISLFSIPVIYERHQVQIDHYLGLANKS 340
 || | :| | || : :
 Db 1184 ----SGAAFRRKRGRQANRKAGCAVTSL-----LASEL 1213
 Qy 341 VK-DAMAKIQAKI 352
 | |||| : ||:
 Db 1214 TKDDAMAAVPAKV 1226

RESULT 15

US-09-323-735-2

; Sequence 2, Application US/09323735

; Patent No. 6197932

; GENERAL INFORMATION:

; APPLICANT: King, Mary-Claire

; APPLICANT: Lynch, Eric D.

; APPLICANT: Lee, Ming

; APPLICANT: Morrow, Jan E.

; APPLICANT: Welcsh, Piri L.

; APPLICANT: Leon, Pedro E.

; TITLE OF INVENTION: Modulators of Actin

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 75 DENISE DRIVE

; CITY: HILLSBOROUGH

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94010

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

```

;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/09/323,735
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  09/080,897
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  OSMAN, RICHARD A
;   REGISTRATION NUMBER:  36,627
;   REFERENCE/DOCKET NUMBER:  UW97-001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (650) 343-4341
;   TELEFAX:  (650) 343-4342
;   INFORMATION FOR SEQ ID NO:  2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  1248 amino acids
;   TYPE:  amino acid
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-09-323-735-2

```

```

Query Match          13.9%;  Score 252.5;  DB 3;  Length 1248;
Best Local Similarity 15.3%;  Pred. No. 0.012;
Matches 140;  Conservative 46;  Mismatches 102;  Indels 625;  Gaps 39;

```

```

Qy      1 MEDIDQ--SSLVSSSTDSPPRP----- 20
      |:| :: |::: || |
Db      378 MDDFNEVFQILLNTVKDSKAEPHFLSILQHLLLVNRNDYEARPQYYKLIEECISQIVLHKN 437

Qy      21 --PPAFKY-----QFVTEPEDEEDEEEE-----E 41
      | || | : : : | : | :
Db      438 GADPDFKCRHLQIEIEGLIDQMIDKTKVEKSEAKAAELEKKLDSELTARHELQVEMKKME 497

Qy      42 EDEEEDDEDLE-----ELEVLE-----RKPAAGLS 66
      | | : :||: | || : :| | ||
Db      498 SDFEQKLQDLQGEKDALHSEKQQIATEKQDLEAEVSQLTGEVAKLTKELEDACKEMASLS 557

Qy      67 AAA--VPPA--AAAPLLDFSSDSVPPAPRGPLPA-----APPAAP--ERQPSWERSPAA 114
      ||| |||: : || |||| || | || | |
Db      558 AAATVPPSVPSRAP-----VPPAP--PLPGDSGTIIPPPAPGDSTTPPPPPPPPP 607

Qy      115 PAPSLPPAAAVLP-----SKLP----- 131
      | | || | : | | | ||
Db      608 PPPPLPGGTALSPPPPLSGDATIPPPPPPLPEGVGIPSPSSLPGGTAIPPPPPPLPGSARIP 667

Qy      132 -----EDDEPPARPP-----PPPPAGASPL 151
      | || || | || | || | :
Db      668 PPPPPLPGSAGIPPPPPPLPGEAGMPPPPPPLPGGPGIPPPPPFPGGPGIPPPPPG---M 724

Qy      152 AEPAAPP---STPAAP-----KRRGSGSSVV--DL---LYWRDIK--- 183
      | || | || | || : | : | || : | : |
Db      725 GMPPPPPFGFGVPAAPVLPFGLTPKKLYKPEVQLRRPNWSKLVAEDLSQDCFWTKVKEDR 784

```

Search completed: September 29, 2004, 18:56:35
Job time : 31.9674 secs

Search completed: September 29, 2004, 18:56:35
Job time : 31.9674 secs

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:54:54 ; Search time 85.8807 Seconds
(without alignments)
1348.937 Million cell updates/sec

Title: SEQ2_1-171_975-1163
Perfect score: 1823
Sequence: 1 MEDIDQSSLVSSSTDSPRP.....VKDAMAKIQAKIPGLKRKAD 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score Match Length DB ID	Description

1	1817	99.7	360	9	US-09-893-348-20	Sequence 20, Appl
2	1808.1	99.2	379	14	US-10-205-194-164	Sequence 164, App
3	1732.7	95.0	1163	9	US-09-893-348-18	Sequence 18, Appl
4	1622.9	89.0	1163	12	US-10-267-502-431	Sequence 431, App
5	1603.3	87.9	373	9	US-09-789-386-6	Sequence 6, Appli
6	1603.3	87.9	373	9	US-09-765-205-6	Sequence 6, Appli
7	1603.3	87.9	373	9	US-09-893-348-24	Sequence 24, Appl
8	1603.3	87.9	373	12	US-10-408-967-8	Sequence 8, Appli
9	1603.3	87.9	373	14	US-10-060-036-72	Sequence 72, Appl
10	1595.3	87.5	373	16	US-10-466-258-4	Sequence 4, Appli
11	1518.4	83.3	1192	9	US-09-789-386-2	Sequence 2, Appli
12	1518.4	83.3	1192	9	US-09-758-140-6	Sequence 6, Appli
13	1518.4	83.3	1192	9	US-09-893-348-23	Sequence 23, Appl
14	1518.4	83.3	1192	9	US-09-972-599A-6	Sequence 6, Appli
15	1518.4	83.3	1192	12	US-10-267-502-429	Sequence 429, App
16	1518.4	83.3	1192	14	US-10-060-036-71	Sequence 71, Appl
17	1518.4	83.3	1192	16	US-10-327-213-9	Sequence 9, Appli
18	1518.4	83.3	1192	16	US-10-466-258-9	Sequence 9, Appli
19	1511.4	82.9	1192	12	US-10-408-967-7	Sequence 7, Appli
20	921	50.5	199	9	US-09-893-348-21	Sequence 21, Appl
21	904	49.6	199	9	US-09-893-348-25	Sequence 25, Appl
22	904	49.6	199	12	US-10-660-946-1	Sequence 1, Appli
23	897	49.2	199	12	US-10-408-967-9	Sequence 9, Appli
24	895	49.1	199	12	US-09-978-360A-467	Sequence 467, App
25	864	47.4	199	16	US-10-466-258-11	Sequence 11, Appl
26	826.3	45.3	776	12	US-10-660-946-5	Sequence 5, Appli
27	826.3	45.3	776	12	US-10-267-502-430	Sequence 430, App
28	823.2	45.2	780	12	US-10-267-502-432	Sequence 432, App
29	807.4	44.3	777	14	US-10-205-219-93	Sequence 93, Appl
30	790	43.3	356	12	US-10-660-946-6	Sequence 6, Appli
31	687.7	37.7	593	15	US-10-108-260A-2892	Sequence 2892, Ap
32	683	37.5	208	12	US-10-660-946-7	Sequence 7, Appli
33	679	37.2	266	12	US-10-276-774-2330	Sequence 2330, Ap
34	672.9	36.9	269	14	US-10-106-698-6222	Sequence 6222, Ap
35	669	36.7	267	12	US-10-660-946-8	Sequence 8, Appli
36	669	36.7	267	14	US-10-205-194-127	Sequence 127, App
37	654.6	35.9	236	9	US-09-729-674-20	Sequence 20, Appl
38	654.6	35.9	236	9	US-09-765-205-26	Sequence 26, Appl
39	654.6	35.9	236	12	US-10-408-967-2	Sequence 2, Appli
40	548.1	30.1	241	12	US-10-660-946-3	Sequence 3, Appli
41	539.9	29.6	168	10	US-09-809-391-563	Sequence 563, App
42	539.9	29.6	168	10	US-09-882-171-563	Sequence 563, App
43	539.9	29.6	168	12	US-10-164-861-563	Sequence 563, App
44	527.7	28.9	222	12	US-10-267-502-428	Sequence 428, App
45	519.6	28.5	234	12	US-10-424-599-200840	Sequence 200840,

ALIGNMENTS

RESULT 1
 US-09-893-348-20
 ; Sequence 20, Application US/09893348
 ; Patent No. US20020072493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EISENBACH-SCHWARTZ, Michal

```

; APPLICANT: COHEN, Irun R.
; APPLICANT: BESERMAN, Pierre
; APPLICANT: MOSONEGO, Alon
; APPLICANT: MOALEM, Gila
; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND
THEIR USES
; FILE REFERENCE: EIS-SCHWARTZ=2A
; CURRENT APPLICATION NUMBER: US/09/893,348
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/314,161
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: US 09/218,277
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: PCT/US98/14715
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: IL 124500
; PRIOR FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-893-348-20

```

```

Query Match          99.7%; Score 1817; DB 9; Length 360;
Best Local Similarity 99.7%; Pred. No. 1.8e-59;
Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60

Qy      61 PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120

Qy      121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPPSTPAAPKRRGSGSSVVDLLYWR 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPPSTPAAPKRRGSGSVVVDLLYWR 180

Qy      181 DIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHP 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 DIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHP 240

Qy      241 FRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 FRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVG 300

Qy      301 ALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360

```

```

RESULT 2
US-10-205-194-164
; Sequence 164, Application US/10205194

```



```

; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; CURRENT FILING DATE: 5200-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 164
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Foocen-m2 reticulon
US-10-205-194-164

```

```

Query Match          99.2%; Score 1808.1; DB 14; Length 379;
Best Local Similarity 94.7%; Pred. No. 4.3e-59;
Matches 359; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

```

```

Qy      1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60
      |
Db      1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60

Qy     61 PAAGLSAAAVPPAAAAPLLDSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120
      |
Db     61 PAAGLSAAAVPPAAAAPLLDSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120

Qy    121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPPSTPAAPKRRGSG----- 170
      |
Db    121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPPSTPAAPKRRGSGSVDETLFALP 180

Qy    171 -----SSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 221
      |:|
Db    181 AASEPVIPISSAVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 240

Qy    222 FRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVD 281
      |
Db    241 FRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVD 300

Qy    282 DLVDSLKFVILMWVFTYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSV 341
      |
Db    301 DLVDSLKFVILMWVFTYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSV 360

Qy    342 KDAMAKIQAKIPGLKRKAD 360
      |
Db    361 KDAMAKIQAKIPGLKRKAD 379

```

RESULT 3

US-09-893-348-18
 ; Sequence 18, Application US/09893348
 ; Patent No. US20020072493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EISENBACH-SCHWARTZ, Michal
 ; APPLICANT: COHEN, Irun R.
 ; APPLICANT: BESERMAN, Pierre
 ; APPLICANT: MOSONEGO, Alon
 ; APPLICANT: MOALEM, Gila
 ; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND
 THEIR USES
 ; FILE REFERENCE: EIS-SCHWARTZ=2A
 ; CURRENT APPLICATION NUMBER: US/09/893,348
 ; CURRENT FILING DATE: 2001-06-28
 ; PRIOR APPLICATION NUMBER: US 09/314,161
 ; PRIOR FILING DATE: 1999-05-19
 ; PRIOR APPLICATION NUMBER: US 09/218,277
 ; PRIOR FILING DATE: 1998-12-22
 ; PRIOR APPLICATION NUMBER: PCT/US98/14715
 ; PRIOR FILING DATE: 1998-07-21
 ; PRIOR APPLICATION NUMBER: IL 124500
 ; PRIOR FILING DATE: 1998-05-19
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 18
 ; LENGTH: 1163
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-09-893-348-18

Query Match 95.0%; Score 1732.7; DB 9; Length 1163;
 Best Local Similarity 31.0%; Pred. No. 2.2e-55;
 Matches 360; Conservative 0; Mismatches 0; Indels 803; Gaps 1;

Qy	1	MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK	60
Db	1	MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK	60
Qy	61	PAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	61	PAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGS-----	171
Db	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETLFALP	180
Qy	172	-----	171
Db	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Qy	172	-----	171
Db	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSEFKGSPKGES	300
Qy	172	-----	171
Db	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQMSVVA	360

Qy	172	-----	171
Db	361	PVREEYADFKPFEQAWVEKDTYEGSRDVLAAARANVESKVDRKCLEDSEQSLGKDSEGR	420
Qy	172	-----	171
Db	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Qy	172	-----	171
Db	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Qy	172	-----	171
Db	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Qy	172	-----	171
Db	601	LPSAGASVVQPSVSPLEAPPPVSYSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Qy	172	-----	171
Db	661	NAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Qy	172	-----	171
Db	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Qy	172	-----	171
Db	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Qy	172	-----	171
Db	841	IIDEFPFTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Qy	172	-----	171
Db	901	DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDTEKED	960
Qy	172	-----SVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	217
Db	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Qy	218	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	277
Db	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Qy	278	FLVDDLVDLKFVAVLMWVFTYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGIA	337
Db	1081	FLVDDLVDLKFVAVLMWVFTYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGIA	1140
Qy	338	NKSVKDAMAKIQAKIPGLKRKAD	360
Db	1141	NKSVKDAMAKIQAKIPGLKRKAD	1163

RESULT 4
 US-10-267-502-431
 ; Sequence 431, Application US/10267502
 ; Publication No. US20040071700A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kim, Jaeseob
 ; APPLICANT: Galant, Ron
 ; TITLE OF INVENTION: Obesity Linked Genes
 ; FILE REFERENCE: LSD-07416
 ; CURRENT APPLICATION NUMBER: US/10/267,502
 ; CURRENT FILING DATE: 2003-01-27
 ; NUMBER OF SEQ ID NOS: 439
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 431
 ; LENGTH: 1163
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-267-502-431

Query Match 89.0%; Score 1622.9; DB 12; Length 1163;
 Best Local Similarity 29.7%; Pred. No. 2.6e-51;
 Matches 347; Conservative 2; Mismatches 7; Indels 811; Gaps 5;

Qy	1	MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK	60
Db	1	MEDIDQSSLVSSSADSPPRRPPPAFKYQFVTEPEDEEDEEDEE-EEEDDEDLEEELEVLERK	59
Qy	61	PAAGLSAAAVPPAAAAPLLDFFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	60	PAAGLSAVPVPP-AAAPLLDFFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP	118
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGS-----	171
Db	119	PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAAPPSTPAAPKRRGSGSVDETIFALP	176
Qy	172	-----	171
Db	177	AASEPVIPSSAEKIMDLKEQPGNTVSSGQEDFPSVLFFETAASLPSLSPLSTVSFKEHGYL	236
Qy	172	-----	171
Db	237	GNLSAVASTEGTIEETLNEASRELPERATNPFVNRESAEFSVLEYSEMGSFFNGSPKGES	296
Qy	172	-----	171
Db	297	AMLVENTKEEVIVRSKDKEDLVCSAALHNPQESPATLTKVVKEDGVMSPKTMDFNEMK	356
Qy	172	-----	171
Db	357	MSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANMESKVDKKCFEDSLEQKSHGK	416
Qy	172	-----	171
Db	417	DSESRNENASFPSTPELVKDGSRAYITCDSFTSATESTAANIFPVLEDHTSENKTDEKKI	476
Qy	172	-----	171

Db	477	EERKAQIITEKTS PKTSNPFLVAIHDSEADYVTTDNLSKVTEAVVATMPEGLTPDLVQEA	536
Qy	172	-----	171
Db	537	CESELNEATGTKIAYETKVDLVQTSEAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEA	596
Qy	172	-----	171
Db	597	PLNSLLPSTGASVAQPSASPLEVPSPVSYDGIKLEPENPPPYEEAMSVALKTSDAKEEIK	656
Qy	172	-----	171
Db	657	EPESFNAAAQEAEPYISIIACDLIKETKLSTEPSPGFSNYSEIAKFEKSVPDHCELVDDS	716
Qy	172	-----	171
Db	717	SPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVSETVTQHKHKERLSASPQEVGKPY	776
Qy	172	-----	171
Db	777	LESFQPNLHITKDAASNEIPTLTKKETISLQMEEFNTAIYSNDDLSSKEDKMKESSETFS	836
Qy	172	-----	171
Db	837	DSSPIEIIDEFPTFVSAKDDSPKEYTDLEVS NKSEIANVQSGANSLPCSELPCDLSFKNT	896
Qy	172	-----	171
Db	897	YPKDEAHVSDEFKSRSSVSKVPLLLPNVSALESQIEMGNIVKPKVLTKEAEKLPDTE	956
Qy	172	-----SVVDLLYWRDIKKTGVV-FGASLFLLLSLTVFSIVSVTAYIAL	213
Db	957	KEDRSLTAVLSAELNKTSSVVDLLYWRDIKKTGVVYFGASLFLLLSLTVFSIVSVTAYIAL	1016
Qy	214	ALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKE	273
Db	1017	ALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKE	1076
Qy	274	LRRFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHY	333
Db	1077	LRRFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHY	1136
Qy	334	LGLANKSVKDAMAKIQAKIPGLKRKAD	360
Db	1137	LGLANKSVKDAMAKIQAKIPGLKRKAE	1163

RESULT 5

US-09-789-386-6

; Sequence 6, Application US/09789386

; Patent No. US20020010324A1

; GENERAL INFORMATION:

; APPLICANT: MICHALOVICH, DAVID

; APPLICANT: PRINJHA, RABINDER KUMAR

; TITLE OF INVENTION: NOVEL COMPOUNDS

; FILE REFERENCE: GP-30165-C1

; CURRENT APPLICATION NUMBER: US/09/789,386
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: U.K. 9916898.1
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: U.K. 9816024.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: US 09/359,208
; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 373
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-789-386-6

Query Match 87.9%; Score 1603.3; DB 9; Length 373;
Best Local Similarity 87.2%; Pred. No. 1.6e-51;
Matches 327; Conservative 11; Mismatches 20; Indels 17; Gaps 6;

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
      |||:||||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPLVSSS-DSPPRQPAPFKYQFVREPEDEE-EEEEEEEEDEDEDLEEELEVLERK 58

Qy      61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || ||| |||:| | :| ||||| |||||:| | :|
Db      59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSTVP 118

Qy      116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPAPAPAAPSTPAAPKR 178

Qy      167 RG-SGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 225
      || ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      179 RGSSGSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 238

Qy      226 KGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVD 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      239 KGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD 298

Qy      286 SLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAM 345
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      299 SLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAM 358

Qy      346 AKIQAKIPGLKRKAD 360
      ||||| ||||| ||||| :
Db      359 AKIQAKIPGLKRKAE 373
```

RESULT 6

US-09-765-205-6

; Sequence 6, Application US/09765205
; Patent No. US20020034800A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Li
; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: 1458.004/200130.449

```
; CURRENT APPLICATION NUMBER: US/09/765,205
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US/09/212,440
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 373
; TYPE: PRT
; ORGANISM: human
US-09-765-205-6
```

```
Query Match          87.9%; Score 1603.3; DB 9; Length 373;
Best Local Similarity 87.2%; Pred. No. 1.6e-51;
Matches 327; Conservative 11; Mismatches 20; Indels 17; Gaps 6;
```

```
Qy      1 MEDIDQSSILVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60
      |||:||||| ||||| ||||| ||||| ||||| |||||:||||| |||||
Db      1 MEDLDQSPLVSSS-DSPRRQPAPFKYQFVREPEDEE-EDEEEEEDEDEDLEEEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || ||| |||:| | :| ||||| ||||| |||||: || : |
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy    116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKR 178

Qy    167 RG-SGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 225
      || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    179 RGSSGSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 238

Qy    226 KGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVD 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    239 KGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD 298

Qy    286 SLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAM 345
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    299 SLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAM 358

Qy    346 AKIQAKIPGLKRKAD 360
      ||||| ||||| ||||| :
Db    359 AKIQAKIPGLKRKAE 373
```

RESULT 7

US-09-893-348-24

; Sequence 24, Application US/09893348

; Patent No. US20020072493A1

; GENERAL INFORMATION:

; APPLICANT: EISENBACH-SCHWARTZ, Michal

; APPLICANT: COHEN, Irun R.

; APPLICANT: BESERMAN, Pierre

; APPLICANT: MOSONEGO, Alon

; APPLICANT: MOALEM, Gila

; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USES

```
; FILE REFERENCE: EIS-SCHWARTZ=2A
; CURRENT APPLICATION NUMBER: US/09/893,348
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/314,161
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: US 09/218,277
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: PCT/US98/14715
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: IL 124500
; PRIOR FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-348-24
```

```
Query Match          87.9%; Score 1603.3; DB 9; Length 373;
Best Local Similarity 87.2%; Pred. No. 1.6e-51;
Matches 327; Conservative 11; Mismatches 20; Indels 17; Gaps 6;
```

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
      |||:||| ||||| ||||| ||||| ||||| |||||:|||:||| ||||| |||||
Db      1 MEDLDQSPLVSSS-DSPPRQPAPFKYQFVREPEDEE-EEEEEEEEDEDEDLEEELEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || || |||:|| |:| ||||| ||||| |||||:|||:||| :| :|
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy    116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKR 178

Qy    167 RG-SGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 225
      || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    179 RGSSGSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 238

Qy    226 KGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVD 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    239 KGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD 298

Qy    286 SLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAM 345
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    299 SLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAM 358

Qy    346 AKIQAKIPGLKRKAD 360
      ||||| ||||| ||||| :
Db    359 AKIQAKIPGLKRKAE 373
```

```
RESULT 8
US-10-408-967-8
; Sequence 8, Application US/10408967
; Publication No. US20040063161A1
; GENERAL INFORMATION:
```


; APPLICANT: Pharmacia & Upjohn Company
 ; APPLICANT: Yan, Riqiang
 ; APPLICANT: Lu, Yifeng
 ; TITLE OF INVENTION: Compositions and Methods of Treating Alzheimer's Disease
 ; FILE REFERENCE: 00925
 ; CURRENT APPLICATION NUMBER: US/10/408,967
 ; CURRENT FILING DATE: 2003-04-08
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 373
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-967-8

Query Match 87.9%; Score 1603.3; DB 12; Length 373;
 Best Local Similarity 87.2%; Pred. No. 1.6e-51;
 Matches 327; Conservative 11; Mismatches 20; Indels 17; Gaps 6;

Qy 1 MEDIDQSSILVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60
 |||:|||| ||||| ||||| ||||| ||||| |||||:||||| |||||
 Db 1 MEDLDQSPLVSSS-DSPPRQPAPFKYQFVREPEDEE-EEEEEEEEDEDEDLEEEVLERK 58
 Qy 61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
 ||||| || ||| |||:| :| ||||| |||||: || : |
 Db 59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVVSSTVP 118
 Qy 116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
 ||| |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPAPAPAAPPSTPAAPKR 178
 Qy 167 RG-SGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 225
 || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 179 RGSSGSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 238
 Qy 226 KGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVD 285
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 239 KGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD 298
 Qy 286 SLKFAVLMWVFTYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAM 345
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 299 SLKFAVLMWVFTYVGALENGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAM 358
 Qy 346 AKIQAKIPGLKRKAD 360
 ||||| ||||| :
 Db 359 AKIQAKIPGLKRKAE 373

RESULT 9
 US-10-060-036-72
 ; Sequence 72, Application US/10060036
 ; Publication No. US20030073144A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Persing, David H.

; APPLICANT: Hepler, William T.
 ; APPLICANT: Jiang, Yuqiu
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
 ; FILE REFERENCE: 210121.566
 ; CURRENT APPLICATION NUMBER: US/10/060,036
 ; CURRENT FILING DATE: 2002-01-30
 ; NUMBER OF SEQ ID NOS: 4560
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 72
 ; LENGTH: 373
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-060-036-72

Query Match 87.9%; Score 1603.3; DB 14; Length 373;
 Best Local Similarity 87.2%; Pred. No. 1.6e-51;
 Matches 327; Conservative 11; Mismatches 20; Indels 17; Gaps 6;

```

Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60
      |||:|||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPVLVSSS-DSPPRQPAPFKYQFVREPEDEE-EEEEEEEEDEDEDLEEEVLERK 58

Qy      61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || ||| ||:| :| ||||| |||||: || : |
Db      59 PAAGLSAAPVFTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy      116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| || || |||||
Db      119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178

Qy      167 RG-SGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 225
      || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      179 RGSSGSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY 238

Qy      226 KGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVD 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      239 KGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD 298

Qy      286 SLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAM 345
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      299 SLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAM 358

Qy      346 AKIQAKIPGLKRKAD 360
      ||||| ||||| :
Db      359 AKIQAKIPGLKRKAE 373
  
```

RESULT 10
 US-10-466-258-4
 ; Sequence 4, Application US/10466258
 ; Publication No. US20040132096A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GLAXO GROUP LIMITED
 ; TITLE OF INVENTION: ASSAY
 ; FILE REFERENCE: P80966 GCW
 ; CURRENT APPLICATION NUMBER: US/10/466,258

```
; CURRENT FILING DATE: 2003-07-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-466-258-4
```

Query Match 87.5%; Score 1595.3; DB 16; Length 373;
Best Local Similarity 86.7%; Pred. No. 3.3e-51;
Matches 325; Conservative 12; Mismatches 21; Indels 17; Gaps 6;

Qy	1	MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK	60
		: : :	
Db	1	MEDLDQSPLVSSS-DSPRRPQPAFKYQFVREPDEE-EEEEEEEEDEDEDLEEEVLERK	58
Qy	61	PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P	115
		: : : :	
Db	59	PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPCWDPSFVSSSTVP	118
Qy	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
		:	
Db	119	APSPLSAAAVSPSKLPQDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR	178
Qy	167	RG-SGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY	225
Db	179	RGSSGSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIY	238
Qy	226	KGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLDV	285
Db	239	KGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLDV	298
Qy	286	SLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAM	345
		: :	
Db	299	SLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAM	358
Qy	346	AKIQAKIPGLKRKAD	360
		:	
Db	359	AKIQAKIPGLKRKAE	373

RESULT 11

US-09-789-386-2

; Sequence 2, Application US/09789386

; Patent No. US20020010324A1

; GENERAL INFORMATION:

; APPLICANT: MICHALOVICH, DAVID

; APPLICANT: PRINJHA, RABINDER KUMAR

; TITLE OF INVENTION: NOVEL COMPOUNDS

; FILE REFERENCE: GP-30165-C1

; CURRENT APPLICATION NUMBER: US/09/789,386

; CURRENT FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: U.K. 9916898.1

; PRIOR FILING DATE: 1999-07-19

; PRIOR APPLICATION NUMBER: U.K. 9816024.5

; PRIOR FILING DATE: 1998-07-22

; PRIOR APPLICATION NUMBER: US 09/359,208
; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-789-386-2

Query Match 83.3%; Score 1518.4; DB 9; Length 1192;
Best Local Similarity 27.3%; Pred. No. 2.1e-47;
Matches 326; Conservative 12; Mismatches 20; Indels 836; Gaps 6;

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60
      |||:||||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPLVSSS-DSPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || || |||:|:|:| ||||| |||||:|:|:|
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSSTVP 118

Qy    116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKR 178

Qy    167 RGSG----- 170
      |||
Db    179 RGSSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSL 238

Qy    171 ----- 170

Db    239 SPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS 298

Qy    171 ----- 170

Db    299 EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVEDE 358

Qy    171 ----- 170

Db    359 VVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLES 418

Qy    171 ----- 170

Db    419 KVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATN 478

Qy    171 ----- 170

Db    479 IFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT KV 538

Qy    171 ----- 170

Db    539 TEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCP 598

Qy    171 ----- 170

Db    599 SFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPP 658
```

Qy	171	-----	170
Db	659	YEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDY	718
Qy	171	-----	170
Db	719	SEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIQDVPQKQDETVMMLVKESLTETSFES	778
Qy	171	-----	170
Db	779	MIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTA	838
Qy	171	-----	170
Db	839	VYSNDDLFIKSEAQIRETETFSDDSPIEIIDEFPTLISSKTDSEFSKLAREYTDLEVSHKS	898
Qy	171	-----	170
Db	899	EIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLLPPDVSA	958
Qy	171	-----SSVVDLLYWRDIKKTG	186
		:	
Db	959	ATQAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTG	1018
Qy	187	VVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLE	246
Db	1019	VVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLE	1078
Qy	247	SEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGL	306
Db	1079	SEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGL	1138
Qy	307	TLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD	360
		: :	
Db	1139	TLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1192

RESULT 12

US-09-758-140-6

; Sequence 6, Application US/09758140

; Patent No. US20020012965A1

; GENERAL INFORMATION:

; APPLICANT: Strittmatter, Stephen M.

; TITLE OF INVENTION: No. US20020012965A1o Receptor-Mediated Blockade of Axonal Growth

; FILE REFERENCE: 44574-5073-US

; CURRENT APPLICATION NUMBER: US/09/758,140

; CURRENT FILING DATE: 2001-01-12

; PRIOR APPLICATION NUMBER: US 60/175,707

; PRIOR FILING DATE: 2000-01-12

; PRIOR APPLICATION NUMBER: US 60/207,366

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 60/236,378

; PRIOR FILING DATE: 2000-09-29

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-140-6

Query Match 83.3%; Score 1518.4; DB 9; Length 1192;
Best Local Similarity 27.3%; Pred. No. 2.1e-47;
Matches 326; Conservative 12; Mismatches 20; Indels 836; Gaps 6;

```
Qy      1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
      |||:||| ||||| ||||| ||||| ||||| ||||| |||||:|||:||| ||||| |||||
Db      1 MEDLDQSPLVSSS-DSPRRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEELEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || || ||| |||:||| :| ||||| ||||| |||||:||| :| :|
Db     59 PAAGLSAAPVPTAPAAGAPIMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVVSSTVP 118

Qy    116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178

Qy    167 RGSG----- 170
      |||
Db    179 RGSSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSL 238

Qy    171 ----- 170
Db    239 SPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS 298

Qy    171 ----- 170
Db    299 EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVKEDE 358

Qy    171 ----- 170
Db    359 VVSSEKAKDSFNEKRVAVEAPMREEYADFKEPFRVWEVKDSKEDSDMLAAGGKIESNLES 418

Qy    171 ----- 170
Db    419 KVDKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATN 478

Qy    171 ----- 170
Db    479 IFPLLGDPTSENKTDEKKIEEKKAIQIVTEKNTSTKTSNPFLLVAAQDSETDYVTTDNLTKV 538

Qy    171 ----- 170
Db    539 TEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCP 598

Qy    171 ----- 170
Db    599 SFESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPP 658

Qy    171 ----- 170
Db    659 YEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIIACDLIKETKLSAEPAPDFSDY 718
```


; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-348-23

Query Match 83.3%; Score 1518.4; DB 9; Length 1192;
Best Local Similarity 27.3%; Pred. No. 2.1e-47;
Matches 326; Conservative 12; Mismatches 20; Indels 836; Gaps 6;

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
      |||:|||| ||||| ||||| ||||| ||||| ||||| |||||:||||| |||||
Db      1 MEDLDQSPVLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEELEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || || ||| |||:| | ||||| ||||| |||||: || : |
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy    116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| || || ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPAPAPAAPPSTPAAPKR 178

Qy    167 RGSG----- 170
      |||
Db    179 RGSSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSL 238

Qy    171 ----- 170

Db    239 SPLSAASFKEHEYLGNLSTVLPTTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS 298

Qy    171 ----- 170

Db    299 EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKEDE 358

Qy    171 ----- 170

Db    359 VVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLES 418

Qy    171 ----- 170

Db    419 KVDKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATN 478

Qy    171 ----- 170

Db    479 IFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT KV 538

Qy    171 ----- 170

Db    539 TEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCP 598

Qy    171 ----- 170

Db    599 SFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPP 658

Qy    171 ----- 170
```


; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-599A-6

Query Match 83.3%; Score 1518.4; DB 9; Length 1192;
Best Local Similarity 27.3%; Pred. No. 2.1e-47;
Matches 326; Conservative 12; Mismatches 20; Indels 836; Gaps 6;

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEDEDLEEEVLERK 60
      |||:|||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEEVLERK 58

Qy      61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || || ||| |||:| | :| ||||| ||||| |||||: || : |
Db      59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy     116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| || || || ||||| |||||
Db     119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178

Qy     167 RGSG----- 170
      |||
Db     179 RGSSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSL 238

Qy     171 ----- 170

Db     239 SPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS 298

Qy     171 ----- 170

Db     299 EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDE 358

Qy     171 ----- 170

Db     359 VVSSEKAKDSFNEKRVAVEAPMREEYADFKEPFEVWEVKDSKEDSDMLAAGGKIESNLES 418

Qy     171 ----- 170

Db     419 KVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATN 478

Qy     171 ----- 170

Db     479 IFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKV 538

Qy     171 ----- 170

Db     539 TEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSSEVMQESLYPAAQLCP 598

Qy     171 ----- 170

Db     599 SFESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPP 658

Qy     171 ----- 170
```

Db	659	YEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDY	718
Qy	171	-----	170
Db	719	SEMAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIQDVPQKQDETVMVLVKESLTETSFES	778
Qy	171	-----	170
Db	779	MIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTA	838
Qy	171	-----	170
Db	839	VYSNDDLFIKSEAQIRETETFSDDSSPIEIIDEFPTLISSKTDSEFSKLAREYTDLEVSHKS	898
Qy	171	-----	170
Db	899	EIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSQVLLLPDVSAL	958
Qy	171	-----SSVVDLLYWRDIKKTG	186
		:	
Db	959	ATQAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTG	1018
Qy	187	VVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLE	246
Db	1019	VVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLE	1078
Qy	247	SEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGFALFNGL	306
Db	1079	SEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGFALFNGL	1138
Qy	307	TLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD	360
		: :	
Db	1139	TLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1192

RESULT 15

US-10-267-502-429
; Sequence 429, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 429
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-429

Query Match 83.3%; Score 1518.4; DB 12; Length 1192;
Best Local Similarity 27.3%; Pred. No. 2.1e-47;
Matches 326; Conservative 12; Mismatches 20; Indels 836; Gaps 6;

QY	1	MEDIDQSSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEDEDLEEELEVLERK	60
Db	1	MEDLDQSPLVSSS-DSPRRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEELEVLERK	58
QY	61	PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P	115
Db	59	PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVPSSTVP	118
QY	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
Db	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR	178
QY	167	RGSG-----	170
Db	179	RGSSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSL	238
QY	171	-----	170
Db	239	SPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS	298
QY	171	-----	170
Db	299	EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVKEDE	358
QY	171	-----	170
Db	359	VVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLES	418
QY	171	-----	170
Db	419	KVDKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATN	478
QY	171	-----	170
Db	479	IFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT KV	538
QY	171	-----	170
Db	539	TEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCP	598
QY	171	-----	170
Db	599	SFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPP	658
QY	171	-----	170
Db	659	YEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIIACDLIKETKLSAEPAPDFSDY	718
QY	171	-----	170
Db	719	SEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI DPV PQQDETVM LVKESLTETSFES	778
QY	171	-----	170
Db	779	MIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELSTA	838

Qy 171 ----- 170

Db 839 VYSNDDLFI SKEAQIRETETFS DSSPIEIIDEFP TLISSKTDSFSK LAREYTDLEVSHKS 898

Qy 171 ----- 170

Db 899 EIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATS KVL LPPDV SAL 958

Qy 171 -----SSVVDLLYWRDIKKTG 186
:|||||||||

Db 959 ATQAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTG 1018

Qy 187 VVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLE 246
|||||||||

Db 1019 VVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLE 1078

Qy 247 SEVAISEELVQKYSNSALGHVNSTIKELRRFLVDDLVDSLKFVILMWVFTYVGALFENGL 306
|||||||||

Db 1079 SEVAISEELVQKYSNSALGHVNCTIKELRRFLVDDLVDSLKFVILMWVFTYVGALFENGL 1138

Qy 307 TLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
|||||||:||||||| |||||:|||||||

Db 1139 TLLILALISLFSVPVIYERHQVQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192

Search completed: September 29, 2004, 19:11:29
Job time : 96.8807 secs

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:43:17 ; Search time 26.4632 Seconds
 (without alignments)
 1308.568 Million cell updates/sec

Title: SEQ2_1-171_975-1163
 Perfect score: 1823
 Sequence: 1 MEDIDQSSSLVSSSTDSPPRP.....VKDAMAKIQAKIPGLKRKAD 360

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.1

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : PIR_78:*
 1: pir1:*
 2: pir2:*
 3: pir3:*
 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	826.3	45.3	776	2	A46583	neuroendocrine-spe
2	683	37.5	208	2	I60904	neuroendocrine-spe
3	669	36.7	267	2	A60021	tropomyosin-relate
4	437.7	24.0	2484	2	T26216	hypothetical prote
5	427.4	23.4	2607	2	T26215	hypothetical prote
6	337.5	18.5	222	2	T26213	hypothetical prote
7	272.1	14.9	786	2	A35466	progesterone recep
8	266.9	14.6	1173	2	T31421	C-terminal domain-
9	265.5	14.6	1611	2	T38236	hypothetical prote
10	258.5	14.2	1058	2	T13286	cappuccino gene pr
11	257.5	14.1	1206	2	S24407	formin isoform IV
12	256.4	14.1	3530	2	A59266	unconventional myo
13	254.9	14.0	1468	2	S11515	formin - mouse

14	254.7	14.0	716	2	T26998	hypothetical prote
15	253.8	13.9	760	2	F86387	probable Pto kinas
16	253.2	13.9	929	2	C96623	hypothetical prote
17	253.2	13.9	1255	2	T31065	diaphanous protein
18	252.2	13.8	2157	2	S71461	proline-rich prote
19	251.9	13.8	1201	2	G86441	unknown protein [i
20	250.9	13.8	1015	2	JC6552	DNA topoisomerase
21	250.2	13.7	907	2	E96636	hypothetical prote
22	246.4	13.5	1657	2	T19536	hypothetical prote
23	246	13.5	1522	2	T39371	transcription regu
24	245.8	13.5	645	2	A71416	hypothetical prote
25	245.4	13.5	1460	1	EDBEIF	immediate-early pr
26	245.2	13.5	3511	2	A59295	unconventional myo
27	245.1	13.4	710	2	D96728	hypothetical prote
28	244.7	13.4	465	2	G02738	FREAC-4 - human
29	244.3	13.4	1064	2	T13963	formin related pro
30	242.7	13.3	1533	2	T00344	hypothetical prote
31	242.1	13.3	1375	2	S48375	hypothetical prote
32	241.9	13.3	1213	2	A58198	serine/proline-ric
33	241.7	13.3	3938	2	T42761	Bassoon protein -
34	241.1	13.2	775	1	EDBE11	immediate-early pr
35	240.9	13.2	980	2	G75523	probable cell divi
36	240.6	13.2	1634	2	T26517	hypothetical prote
37	240.5	13.2	1446	1	A45344	immediate-early pr
38	238.9	13.1	731	2	T04455	hypothetical prote
39	238.5	13.1	1047	2	A55617	masquerade precurs
40	238.5	13.1	1420	2	T37781	probable cytoskele
41	238.4	13.1	7962	2	I38346	elastic titin - hu
42	238	13.1	3149	1	QQBE8	BPLF1 protein - hu
43	237.6	13.0	933	1	QRHUP	progesterone recep
44	237.5	13.0	980	2	S54986	regulatory protein
45	236.7	13.0	1870	2	S37671	MHC class III hist

ALIGNMENTS

RESULT 1

A46583

neuroendocrine-specific protein, splice form A - human

N;Contains: neuroendocrine-specific protein, splice form B

C;Species: Homo sapiens (man)

C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 05-Nov-1999

C;Accession: A46583; I60903

R;Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers, F.C.; Van de Ven, W.J.

J. Biol. Chem. 268, 13439-13447, 1993

A;Title: Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product.

A;Reference number: A46583; MUID:93293865; PMID:7685762

A;Accession: A46583

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-776 <ROE1>

A;Cross-references: GB:L10333; NID:g307306; PIDN:AAA59950.1; PID:g307307

A;Accession: I60903

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 421-776 <ROE2>
A;Cross-references: GB:L10334; NID:g307308; PIDN:AAA59951.1; PID:g307309
C;Genetics:
A;Gene: GDB:RTN1; NSP
A;Cross-references: GDB:203968; OMIM:600865
A;Map position: 14q21-14q22

Query Match 45.3%; Score 826.3; DB 2; Length 776;
Best Local Similarity 27.7%; Pred. No. 3.9e-20;
Matches 195; Conservative 52; Mismatches 91; Indels 367; Gaps 17;

```

Qy      10 VSSSTDs-----PPR----- 19
      |||: |
Db      85 VSSAMDHTFSTTSKDGEGSCYTSLSIDICYPPQEDSTYFTGILQKENGHVITISESPEELG 144

Qy      20 -PPP-----AFKYQFVTEPED 34
      | |
Db     145 TPGPSLPDVPGIESRGLFSSDSGIEMTPAESTEVNKKILADPLDQMKAEAYKYIDITRPEE 204

Qy      35 EEDEEEEEDEEED-----DED----- 50
      : :| : | | |
Db     205 VKHQEQHHPELEDKDLDFKNKDTDISIKPEGVREPDKPAPVEGKIIKDHLLLEESTFAPYI 264

Qy      51 -----LEELE----- 55
      | | :|
Db     265 DDLSEEQRRAPQITTPVKITLTEIEPSVETTTQEKTPKQDICKPSPDVPTVTVTSEPE 324

Qy      56 -----VLERKPAAGLS----- 66
      : | | | |
Db     325 DDSPGSITPPSSGTEPSAAESQKGKSISEDELITAIKEAKGLSYETAENPRPVGQLADRP 384

Qy      67 ---AAAVPPAAAAPLLDFSSDSV-----PPA 89
      | : | | : | | : |
Db     385 EVKARSGPPTIPSP-LDHEASSAESGDSEIELVSEDPMMAEDALPSGYVSFGHVGGPPPS 443

Qy      90 PRGPL-----PAAPPA----- 100
      | |
Db     444 PASPSIQYSILREEREAEELDSIELIESCDASSASEESPKREQDSPMKPSALDAIREETG 503

Qy     101 --APERQPS-----WERSPAAPAPSLPPA-----AAVLPSKLPEDD----E 135
      | | | | : : | | | | : | | | | :
Db     504 VRAEERAPSRRLAEPGSFLDYPSTEPQGPPELPPGDGALEPETPMLPRK-PEEDSSSNQ 562

Qy     136 PPARPPPPPPAGASPLAEPAPPPSTPAAPKRRGSGSSVVDLLYWRDIKKTGVVFGASLFL 195
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     563 SPAATKGGPLG-----PGAPPPLFLNKQK-----AIDLLYWRDIKQTGIVFGSFLLL 611

Qy     196 LLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEEL 255
      | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db     612 LFSLTQFSVVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQ 671

Qy     256 VQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALIS 315
      : | | : : : | | | | | | | | | | | | | | | | | | | | | |
Db     672 IQKYTDCLQFYVNSTLTKELRRLFLVQDLVDLSLKFAVLMWLLTYVGALFNGLTLLLMAVVS 731

```


Qy 316 LFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRRKAD 360
 :|::||:| :|| ||| |||| : :||| ||||| || |:
 Db 732 MFTLPVVYVKHQAQIDQYLGLVRTHINAVVAKIQAKIPGAKRHAE 776

RESULT 2

I60904

neuroendocrine-specific protein C - human

C;Species: Homo sapiens (man)

C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 05-Nov-1999

C;Accession: I60904

R;Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers, F.C.; Van de Ven, W.J.

J. Biol. Chem. 268, 13439-13447, 1993

A;Title: Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product.

A;Reference number: A46583; MUID:93293865; PMID:7685762

A;Accession: I60904

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-208 <RES>

A;Cross-references: GB:L10335; NID:g307310; PIDN:AAA59952.1; PID:g307311

C;Genetics:

A;Gene: GDB:RTN1; NSP

A;Cross-references: GDB:203968; OMIM:600865

A;Map position: 14q21-14q22

Query Match 37.5%; Score 683; DB 2; Length 208;
 Best Local Similarity 67.4%; Pred. No. 1.1e-16;
 Matches 128; Conservative 31; Mismatches 31; Indels 0; Gaps 0;

Qy 171 SSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQ 230
 | :||| ||||| :||: ||| | || ||| ||: ||| ||: ||| || ||| ||||| | :|
 Db 19 SQAIDLLYWRDIKQTGIVFGSFLLLLSLTQFSVSVVAYLALAALSATISFRIYKSVLQ 78

Qy 231 AIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFA 290
 |: ||: ||||| :||| | : :| : ||| : : ||| : ||||| ||||| ||||| |||||
 Db 79 AVQKTDEGHPFKAYLELEITLSQEIQIKYTDCLQFYVNSTLKELRRLFLVQDLVDLSLKFA 138

Qy 291 VLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQA 350
 ||||: ||||| ||||| :||: ||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 139 VLMWLLTYVGALFNGLTLLLMAVVSMTLPVVYVKHQAQIDQYLGLVRTHINAVVAKIQA 198

Qy 351 KIPGLKRRKAD 360
 |||| || |:
 Db 199 KIPGAKRHAE 208

RESULT 3

A60021

tropomyosin-related protein, neuronal - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 05-Nov-1999

C;Accession: A60021

R;Wieczorek, D.F.; Hughes, S.R.

Brain Res. Mol. Brain Res. 10, 33-41, 1991

A;Title: Developmentally regulated cDNA expressed exclusively in neural tissue.
 A;Reference number: A60021; MUID:91278684; PMID:1647480
 A;Accession: A60021
 A;Molecule type: mRNA
 A;Residues: 1-267 <WIE>
 A;Cross-references: EMBL:X52817; NID:g456549; PIDN:CAA37001.1; PID:g456550
 C;Comment: This neuronal-specific mRNA was identified by hybridization to an alpha-tropomyosin probe but does not show homology in amino acid sequence.

Query Match 36.7%; Score 669; DB 2; Length 267;
 Best Local Similarity 66.7%; Pred. No. 6.4e-16;
 Matches 124; Conservative 32; Mismatches 30; Indels 0; Gaps 0;

```

Qy      171 SSVVDLLYWDRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQ 230
      | :|||||||:|:|:| | | | | | :| | | | | | | | | | | | | | | | | |
Db      10 SQAIDLLYWDRDIKQTGIVFGSFLLLLSLTQFSVSVVAYLALAALSATISFRIYKSVLQ 69

Qy      231 AIQKSDEGHPPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDSLKFA 290
      |:|:|:|:|:|:|:|:| | | :|:| | | :|:|:| :|:|:|:|:|:|:|:|:|:|
Db      70 AVQKTDEGHPPFKAYLELEITLSQEIQKYTDCLQLYVNSTLKELRRLFLVQDLVDSLKFA 129

Qy      291 VLMWVFETYVGALENGLTLLILALISLSFIPVIYERHQVQIDHYLGLANKSVKDAMAKIQA 350
      |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      130 VLMWLLTYVGALENGLTLLMAVVSMTLPVVYVKHQAQVDQYLGLVRTHINTVVAKIQA 189

Qy      351 KIPGLK 356
      ||| | :
Db      190 KIPGAR 195
  
```

RESULT 4

T26216

hypothetical protein W06A7.3c - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T26216

R;Ainscough, R.

submitted to the EMBL Data Library, August 1996

A;Reference number: Z20173

A;Accession: T26216

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-2484 <WIL>

A;Cross-references: EMBL:Z78066; PIDN:CAB51467.1; GSPDB:GN00023; CESP:W06A7.3c

A;Experimental source: clone W06A7

C;Genetics:

A;Gene: CESP:W06A7.3c

A;Map position: 5

A;Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2339/2; 2463/2

Query Match 24.0%; Score 437.7; DB 2; Length 2484;
 Best Local Similarity 14.8%; Pred. No. 1.2e-05;
 Matches 135; Conservative 77; Mismatches 138; Indels 563; Gaps 21;

```

Qy      1 MEDIDQSSL-----VSSST----- 14
      || : : | : || ||
Db      1566 MEVVTSEISEMAPQVSESTCPIPEPLADLKLPEVDEKTPPEPEPVVPGQVQERIPIEV 1625
  
```

Qy 15 ----DSPRRPPPAFKYQF--VTEPEDE----- 35
| | | | | : | : | :
Db 1626 EQAPTIPQRPPRAPKSELPKVAKPLDDSKSRVRFAPLNIKLGRTYSEEQQKELVESLERP 1685

Qy 36 -----ED-----EEEEDE----- 44
| | | : | | |
Db 1686 LTIITQQKPPEKPTEDIGALSPLSPNTLAEYEEVPMMDMQSVPHSPQEKQEEIEALSEII 1745

Qy 45 -----EEDDEDLEELEVL----- 57
| : | | | | : :
Db 1746 EEPQAMKEVEKPVESAPEKDNESLEAPEIINEPIRRVLVETKIMGPGKSLNEDNDDDDDG 1805

Qy 58 ----- 57
Db 1806 SECLDSIGDLSSERTIQRFNTSIDDPSIRRDSFSSISSFGDRQKFRTAIENIRQDLLPFQS 1865

Qy 58 -----ERKPAA----- 63
| : | : |
Db 1866 SVSQYLRRSSPNPSQQLLVTNLSMDSPSDLSPNAPPVGFENTAQFLEKLQQEDRPSAEGSI 1925

Qy 64 -----GLSAAAVPP----- 72
| | | | |
Db 1926 DSSGFEEKVDHEGLDEFAAPPVHDPMQKSVFGSLGSDDMKPGSQDDGFVFIERNEANEATL 1985

Qy 73 -----AAAAPLL----- 79
| | | | |
Db 1986 KKNQKMSSHHNDVIEKNYFNDNAPTAALLESPIAEEARLKVQDAVESASEYKKQAVDSGD 2045

Qy 80 -----DFSSDSVPPA-----PRGPL 94
| | : | | | | :
Db 2046 EIGRELLDNVEQKIEQVKEPIVDSLHKAYDGVGDFVHETVPNAVDDFVREAQKQLPESPV 2105

Qy 95 P-----AAPP-----AAP-----ERQ 105
| | | | | :
Db 2106 PEKIETPEPLVDIHDTVDKVHDEVDFNLRREPTPPFETDDVAPLSDDKPQFGNQTPPEDE 2165

Qy 106 PSWER-----SPAAPAPSLPPAAAVLPSKLPEDD 134
: : | : | | : : | | : :
Db 2166 TTFDRKGPLTIPEEVEKAAAAQNNDLDDFDPLVTSNTGAAFGAAVGAAA--ESLTEEE 2223

Qy 135 -----EPPARPPPPPPAGASPLAEPAAPPST-----PAAP-----KRRGSG- 170
| | | | | : : | | | : : | | |
Db 2224 MFGHQKFETVPRPPTPP---KDISEDVKPSTVNLGPSHHHSHSPSPHHSILKHHGDAW 2279

Qy 171 -----SSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFR 223
| : : | | | | : : | : | : : : | | | |
Db 2280 IDEKTVPPCVLDVIYWRDAKSAIVLSLALLVLFVLAKYPLLTVVVTSLLLALGAAAGFR 2339

Qy 224 IYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDL 283
: : | | : : | | : : : : | : : : | : : : | : :
Db 2340 VFVKVEAQIKKTDSEHPFSEILAQDLTLPOQEKVHAQADVFEHATCIANKLKKLVFVESP 2399

Qy 284 VDSLKFVAVLMWVFTYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKD 343
: : | | : : | | : : | | | | : : | | : : | : : : : :
Db 2400 LESIKFGLVLWSLTYIASWFSGFTLAILGLLGVSVPKVYESNQEAIDPHLATISGHLKN 2459

Qy 344 AMAKIQAKIPGLK 356
| | : | | :
Db 2460 VQNIIDEKLPFLR 2472

RESULT 5

T26215

hypothetical protein W06A7.3a - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T26215

R;Ainscough, R.

submitted to the EMBL Data Library, August 1996

A;Reference number: Z20173

A;Accession: T26215

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-2607 <WIL>

A;Cross-references: EMBL:Z78066; PIDN:CAB01522.2; GSPDB:GN00023; CESP:W06A7.3a

A;Experimental source: clone W06A7

C;Genetics:

A;Gene: CESP:W06A7.3a

A;Map position: 5

A;Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2412/1; 2462/2;
2586/2

Query Match 23.4%; Score 427.4; DB 2; Length 2607;
Best Local Similarity 13.0%; Pred. No. 3.1e-05;
Matches 135; Conservative 78; Mismatches 137; Indels 686; Gaps 21;

```
Qy      1 MEDIDQSSL-----VSSST----- 14
      || : | : || ||
Db    1566 MEVVTSEISEMAPQVSESTCPIPEPLADLKL PVEDDEKTPEPEPVVPGQVQERIIPIEV 1625

Qy      15 ----DSPPRPPPAFKYQF--VTEPEDE----- 35
      | ||| | | : | : | | :
Db    1626 EQAPTIPQRPPRAPKSEL PKVAKPLDDSKSRVRFAPLNIKLGRTYSEEQQKELVESLERP 1685

Qy      36 -----ED-----EEEEDE----- 44
      || || : || |
Db    1686 LTIITQQKPPEKPTEDIGALSPLSPNTLA EYEEVPMMDMQSVPHSPQEKQEEIEALSEII 1745

Qy      45 -----EEDDEDLEELEVL----- 57
      | : | : | || | : :
Db    1746 EEPQAMKEVEKPVESAPEKDNESLEAPEI INEPIRRVLVETKIMGPGKSLNEDNDDDDDG 1805

Qy      58 ----- 57

Db    1806 SECLDSIGDLSSERTIQRFNTSIDDPSIR RDSFSSISSFGDRQKFRTAIENIRQDLLPFQS 1865

Qy      58 -----ERKPAA----- 63
      | : | : |
Db    1866 SVSQYLRSSPNPSQQLLVTNLSMDSPS DLSPNAPPVGFENTAQFLEKLQQEDRPSAEGSI 1925

Qy      64 -----GLSAAAVPP----- 72
      || | ||
Db    1926 DSSGFEEKVDHEGLDEFAAPPVHDPMQKSVFGSLGSDDMKPGSQDDGFVFIERNEANEATL 1985
```

Qy 73 -----AAAAPLL----- 79
 | | | |
 Db 1986 KKNQKMSSHHNDVIEKNYFNDNAPTAALLESPIAEEARKLVQDAVESASEYKKQAVDSGD 2045

Qy 80 -----DFSSDSVPPA-----PRGPL 94
 | | : | | | | :
 Db 2046 EIGRELLDNVEQKIEQVKEPIVDSLHKAYDGVGDFVHETVPNAVDDFVREAQKQLPESPV 2105

Qy 95 P-----AAPP-----AAP-----ERQ 105
 | | | |
 Db 2106 PEKIETPEPLVDIHD TVDKVHDEVNDFLRREPTPPFETDDVAPLSDDKPQFGNQTPPEDE 2165

Qy 106 PSWER-----SPAAPAPSLPPAAAVLPSKLPEDD 134
 : : : | : | | : : | | | | | : :
 Db 2166 TTFDRKGPLTIPEEVEKAAAAQNNDLDDFDPLVTSNTGAAFGAAVGAAGAAV--ESLTEEE 2223

Qy 135 -----EPPARPPPPPPAGASPLAEPAPST-----PAAPKR----- 166
 | | | | : : | | | : :
 Db 2224 MFGHQKFETVPRPPTPP----KDISDEDVKPSTVNLGPSHHSHSPSSPHHSILKHHGDAW 2279

Qy 167 ----- 166

Db 2280 IDFKTVPPCAQNAFSPGEIMFLLAFFVYLSCFASFFSKSLPLLDNLLSLVYLSISLIH 2339

Qy 167 ----- 166

Db 2340 VKHHRKFRWNEEQATTMSKLGAVGRGLYALIAFIVNIVLRVGLNVALVVGVAVSAHEAYK 2399

Qy 167 --RSGS----SSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTI 220
 : | | | : : | | | : : | : : : | | |
 Db 2400 LTKSSGVLRRKKEVLDVIYWRDAKKSAILVSLALLVLFVLAKYPLLTVVVTYSLLIALGAAA 2459

Qy 221 SFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLV 280
 | : : | | : : | | | : : : | | : : | | : : | |
 Db 2460 GFRVFKKVEAQIKKTDSEHPFSEILAQDLTLPOEKVHAQADVFEHATCIANKLKKLVFV 2519

Qy 281 DDLVDSLKFAVIMWVFTYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKS 340
 : : : | | : : | | : : | | | : : | | : : | | : :
 Db 2520 ESPLESIKFGVLVLSLTYIASWFSGFTLAILGLLVFVSVPKVYESNQEAIDPHLATISGH 2579

Qy 341 VKDAMAKIQAKIPGLK 356
 : | : | : | :
 Db 2580 LKNVQNIIDEKLPFLR 2595

RESULT 6

T26213

hypothetical protein W06A7.3b - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T26213

R;Ainscough, R.

submitted to the EMBL Data Library, August 1996

A;Reference number: Z20173

A;Accession: T26213

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA
 A;Residues: 1-222 <WIL>
 A;Cross-references: EMBL:Z78066; PIDN:CAB01523.1; GSPDB:GN00023; CESP:W06A7.3b
 A;Experimental source: clone W06A7
 C;Genetics:
 A;Gene: CESP:W06A7.3b
 A;Map position: 5
 A;Introns: 27/1; 77/2; 201/2

Query Match 18.5%; Score 337.5; DB 2; Length 222;
 Best Local Similarity 25.7%; Pred. No. 8.1e-05;
 Matches 65; Conservative 49; Mismatches 84; Indels 55; Gaps 2;

```

Qy      104 RQPSWERSPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPSTPAA 163
          :||:|      ||  |
Db      13 KQPTW-----VPATDFP----- 24

Qy      164 PKRRGSGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFR 223
          ::||:|||| ||: :| :| :| | : ::| | | | ||
Db      25 -----GKILDVIYWRDAKKSIAIVLSIALLVLFVLAKYPLLTVVVTSLLLALGAAAGFR 77

Qy      224 IYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDL 283
          ::| | | :||:| ||| | :: : :| | : : | :| :| | :
Db      78 VFKKVEAQIKKTDSEHPFSEILAQDLTLPQEKVHAQADVFEHATCIANKLKKLVFVESP 137

Qy      284 VDSLKFAVLMWVFETYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKD 343
          ::||:| ::| ||: : | :| || || | : ||:| :|| :| || :| : :|
Db      138 LESIKFGLVLWSLTYIASWFSGFTLAILGLLGVSVPKVYESNQEAIDPHLATISGHLKN 197

Qy      344 AMAKIQAKIPGLK 356
          | | :| | :
Db      198 VQNIIDEKLPFLR 210
  
```

RESULT 7

A35466
 progesterone receptor form B - chicken
 N;Contains: progesterone receptor form A
 C;Species: Gallus gallus (chicken)
 C;Date: 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
 C;Accession: A35466; S06284; A40903; A24661; A24312; A40911; A61552
 R;Jeltsch, J.M.; Turcotte, B.; Garnier, J.M.; Lerouge, T.; Krozowski, Z.; Gronemeyer, H.; Chambon, P.
 J. Biol. Chem. 265, 3967-3974, 1990
 A;Title: Characterization of multiple mRNAs originating from the chicken progesterone receptor gene. Evidence for a specific transcript encoding form A.
 A;Reference number: A35466; MUID:90154085; PMID:2303488
 A;Accession: A35466
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-786 <JEL>
 A;Cross-references: GB:M32732; GB:J05240; NID:g212554; PIDN:AAA49011.1; PID:g212558
 R;Gronemeyer, H.; Turcotte, B.; Quirin-Stricker, C.; Bocquel, M.T.; Meyer, M.E.; Krozowski, Z.; Jeltsch, J.M.; Lerouge, T.; Garnier, J.M.; Chambon, P.
 EMBO J. 6, 3985-3994, 1987

A;Title: The chicken progesterone receptor: sequence, expression and functional analysis.
 A;Reference number: S06284; MUID:88166640; PMID:3443098
 A;Accession: S06284
 A;Molecule type: DNA
 A;Residues: 1-786 <GRO>
 A;Cross-references: EMBL:Y00092; NID:g63744; PIDN:CAA68282.1; PID:g63745
 R;Conneely, O.M.; Dobson, A.D.W.; Tsai, M.J.; Beattie, W.G.; Toft, D.O.; Huckaby, C.S.; Zarucki, T.; Schrader, W.T.; O'Malley, B.W.
 Mol. Endocrinol. 1, 517-525, 1987
 A;Title: Sequence and expression of a functional chicken progesterone receptor.
 A;Reference number: A40903; MUID:91042592; PMID:3153474
 A;Accession: A40903
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-64,'E',65-786 <CON>
 A;Cross-references: GB:M37518
 R;Conneely, O.M.; Sullivan, W.P.; Toft, D.O.; Birnbaumer, M.; Cook, R.G.; Maxwell, B.L.; Zarucki-Schulz, T.; Greene, G.L.; Schrader, W.T.; O'Malley, B.W.
 Science 233, 767-770, 1986
 A;Title: Molecular cloning of the chicken progesterone receptor.
 A;Reference number: A24661; MUID:86289413; PMID:2426779
 A;Accession: A24661
 A;Molecule type: mRNA
 A;Residues: 128-133,'E',135-147,'E',149-164 <CO2>
 A;Note: the authors translated the codon CAG for residue 7 as Glu and CAG for residue 21 as Glu
 R;Jeltsch, J.M.; Krozowski, Z.; Quirin-Stricker, C.; Gronemeyer, H.; Simpson, R.J.; Garnier, J.M.; Krust, A.; Jacob, F.; Chambon, P.
 Proc. Natl. Acad. Sci. U.S.A. 83, 5424-5428, 1986
 A;Title: Cloning of the chicken progesterone receptor.
 A;Reference number: A24312; MUID:86287271; PMID:2426697
 A;Accession: A24312
 A;Molecule type: mRNA
 A;Residues: 417-490 <JE2>
 A;Cross-references: GB:M14280; NID:g212607; PIDN:AAA49039.1; PID:g212608
 A;Note: amino acid and corresponding nucleotide sequences are also shown for three smaller peptides
 R;Birnbaumer, M.; Hinrichs-Rosello, M.V.; Cook, R.G.; Schrader, W.T.; O'Malley, B.W.
 Mol. Endocrinol. 1, 249-259, 1987
 A;Title: Chemical and antigenic properties of pure 108,000 molecular weight chick progesterone receptor.
 A;Reference number: A40911; MUID:88288199; PMID:3453892
 A;Accession: A40911
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 128-133,'E',135-147,'E',149-164;546-558 <BIR>
 R;Simpson, R.J.; Grego, B.; Govindan, M.V.; Gronemeyer, H.
 Mol. Cell. Endocrinol. 52, 177-184, 1987
 A;Title: Peptide sequencing of the chick oviduct progesterone receptor form B.
 A;Reference number: A61552; MUID:88005426; PMID:3653503
 A;Accession: A61552
 A;Molecule type: protein
 A;Residues: 136-153;168-174;195-228;526-537,'X',539;546-563 <SIM>
 C;Genetics:
 A;Introns: 400/2; 451/1; 490/1; 591/1; 639/2; 683/1; 735/3

C;Superfamily: progesterone receptor; erbA transforming protein homology
 C;Keywords: DNA binding; nucleus; steroid hormone receptor; zinc finger
 F;1-786/Product: progesterone receptor form B #status predicted <MA1>
 F;128-786/Product: progesterone receptor form A #status predicted <MA2>
 F;419-682/Domain: erbA transforming protein homology <ERBA>
 F;421-441/Region: zinc finger
 F;457-481/Region: zinc finger

Query Match 14.9%; Score 272.1; DB 2; Length 786;
 Best Local Similarity 15.9%; Pred. No. 0.31;
 Matches 121; Conservative 54; Mismatches 119; Indels 469; Gaps 28;

```

Qy      11 SSSTDSP-----RPPA-----FKYQFVTEPEDEDEDEDEDEEDD 48
      | | : |           : ||:           | : | : | : | | : | | :
Db      6 SKETRAPSSARDGAVLLQAPPSRGEAEGIDVALDGLLYPRSSDEEEEEEEENEEEEEEEP 65

Qy     49 EDLEELEVLER-----KPAAG-----LSAAAVPPAAAAAPLLDFSSD--SVPP 88
      : || | | |           : | |           | | | | | | | : ||
Db     66 QQREEEEEEEEEEDRDCPSYRPGGSLSKDCLDSVLDLFLAPAAHAAPWSLFGPEVPEVPV 125

Qy     89 AP--RGP----LPAAPPAAPERQPSWERSPAAP-----AP 117
      || || | : | | | || | | | || | | | |
Db    126 APMSRGPEQKAVDAGPGAPGPSQP----RPGAPLWPGADSLNVAVKARPGPEDASENRAP 181

Qy    118 SLP-----PAAAVLPSKLPEDDEP-----136
      || | | | | | | | | | | | | | | | |
Db    182 GLPGAEEERGFPERDAGPGEGGLAPAAAASPAAV---EPGAGQDYLVHPILPLNSAFLAS 237

Qy    137 -----136

Db    238 RTRQLLDVEAAYDGSFAFGPRSSPSVPAADLAEGYGYPPPDGKEGPFAYGEFQSALKIKEEG 297

Qy    137 ---PARPPP-----PPPAGASPLAE-----PAAP---PST 160
      || || | | | | | | | | | | | | | | | |
Db    298 VGLPAAPPPFLGAKAAPADFAQPPRAGQEPSLECVLYKAEPPLLPGAYGPPAAPDSL PST 357

Qy    161 PAAP-----164
      || |
Db    358 SAAPPGLYSPLGLNGHHQALGFPAAVLKEGLPQLCPPYLG YVRPDTETSQSSQYSFESLP 417

Qy    165 -----KR-----RSGSSVVDLL-----Y 178
      | | | | | | | | | | | | | | | |
Db    418 QKICLICGDEASGCHYGVLTCGCKVFFKRAMEGQHNYLCAGRNDICVDKIRRNKCPACR 477

Qy    179 WRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEG 238
      | : | : | | | | : | | | | | | | | | |
Db    478 LRKCCQAGMVLGGRKFKKLN-----KMKVVRTL DVALQQP---512

Qy    239 HPFRAYLESEVAISEELVQKYS---NSALGHVNSTIKELR-----275
      | | : | : | | : | | : | | | |
Db    513 ----AVLQDE---TQSLTQRLSFSPNQEI PFVPPMISVLRGIEPEVVYAGYDNTK PETPS 565

Qy    276 -----RLFLVDDLVDLKFVAVLMWVFTYVGALFNGLTL 308
      | : | | : : : : | | | | |
Db    566 SLLTSLNHL CERQLLCVVKWSKLLPGFRNLHIDDQITLIQYS---W-----MSL 611

Qy    309 LILAL-----ISLFSIPVIYERHQVQIDHY 333

```



```

      :: |:                               :|:: :| : | || : :
Db      612 MVFAMGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLSMWQLPQEFVRLQVSQEEF 671
Qy      334 L-----GLANKSVKDAM----AKIQAKIPGLKRRK 358
      |               || ::| | | : | ||::|
Db      672 LCMKALLLLNTIPLEGLRSQSQFDEMRTSYIRELVKAIGLRQK 714

```

RESULT 8

T31421

C-terminal domain-binding protein rA1 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 07-Dec-1999

C;Accession: T31421

R;Yuryev, A.; Patturajan, M.; Litington, Y.; Joshi, R.V.; Gentile, C.; Gebara, M.; Corden, J.L.

Proc. Natl. Acad. Sci. U.S.A. 93, 6975-6980, 1996

A;Title: The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins.

A;Reference number: Z21024; MUID:96293459; PMID:8692929

A;Accession: T31421

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1173 <YUR>

A;Cross-references: EMBL:U49056; NID:g1438531; PID:g1438532; PIDN:AAC52657.1

A;Experimental source: hippocampus

```

Query Match          14.6%;  Score 266.9;  DB 2;  Length 1173;
Best Local Similarity 9.8%;  Pred. No. 1.3;
Matches 116;  Conservative 32;  Mismatches 69;  Indels 961;  Gaps 26;

```

```

Qy      11 SSSTDSPPRPPP-----AF-----KYQ----- 27
      |||: ||| |||               |:         ||
Db      119 SSSSPSPPPPPPPPPPPALPAPRFDIYDPFHPTDEAYSPPPAPEQKYDPFEATGSNPSSS 178
Qy      28 -FVTEPEDEEDEEEEEDEE----- 45
      ||: ||: ||| ||: ||
Db      179 GGTPSPEEEEEEEEEEEEEGLSQSIRRISETLAGIYDDNSLSQDFPGDDSPHREPPPPQT 238
Qy      46 -----EDD 48
      |
Db      239 LGAPGTPPQADSTRAEGAPRRRVFVVGPEAEACLEGKVSVEVVTAGGPALPLPPLPPTD 298
Qy      49 EDLEELEVLE-----RKP-----AAGLSAAAVPPAAA--AP----LL 79
      ::|| |:::               |:|         | ::| | ||:| || |
Db      299 PEIEEGEIVQPEEEPRVAVSLFRAARPRQPPASVATLASVAAPAAPPASAPRAPEGDDFL 358
Qy      80 DFSSDS-----VPPAP-----RGPLPAAPPA-- 100
      :||               ||||               | | |||
Db      359 SLHADSDGEGALQVDLGEPAPPAADARWGGLDLRRKILTQRRERYRQRSASPGPPPPARK 418
Qy      101 -----APERQPSWE----- 109
      ||   |:|
Db      419 KARRERQSRGDPAPPDSPTWEAKKHSRERKLGSHSTARRRSRSRRRRSRSDARRRG 478
Qy      110 -----RS-----PAA----- 114
      ||   |||

```

Db	479	SHRSRSREKRRRRRRSASPPPAASSSSSSRRERHRGKRREGGKKKKKRSRSRAEKRSGL	538
Qy	115	---PAPSLP-----PAAAV	125
Db	539	EKLPAVPVPSGSDRDSRRRGAVPPSIQDLTDHDLFAIKRTITVGRPKTEPRAPSPAPAV	598
Qy	126	LPS-----	128
Db	599	SPKGEVLYDSEGLSADERGAKGDKDRRRSGAASSSSSSREKASRRKALDGDRGRDRDRSS	658
Qy	129	---KLPEDDEP-----	136
		: :	
Db	659	KKPRTPKDSAPGSGALPKAPPRSGSSSSSSSSCSSRKVKLQSKVAVLIREGVSSSTTPAKDS	718
Qy	137	-----	136
Db	719	SSSGLGSIGVKFSRDRESRSPFLKPDERSPAEGVKVAPGSTKPKKTKAKAKAGAKKAKGT	778
Qy	137	-----P	137
Db	779	KGKTKPSKTRKKVRSRSGSSTASGGPGSLKSKADSCSQAASAKGTEETSWSGEERTTKAP	838
Qy	138	ARPP---PPPP-----	145
		:	
Db	839	STPPPKVAPPPPALTTPDSQTVDSCKTPDVSLAEASEDTGVRVGAEEEEEEEEEEEEEE	898
Qy	146	-----AGASPLAEPAA-----	156
		: :	
Db	899	EEQQPATTATSTAAAPSTAPSAGSTAGDGAEDGPAARASQLPTLPPMPWNLPAVD	958
Qy	157	-----PPST-----	160
Db	959	CTTSGVLALTALLFKMEEANLASRAKAQELIQDTNQILRHRKPPSTLGVTPAPVPTSFGL	1018
Qy	161	-----PAAPKRRSGSSVVDLLYWRDIKKTGVVFGA	191
		:	
Db	1019	PPAPSSYLLPGSLPIGGCGSTPPTPTGLVPASDKREGSSSS-----	1059
Qy	192	SLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAI	251
Db	1060	-----EGR-----	1062
Qy	252	SEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLIL	311
		: : :	
Db	1063	-----GDTDKYLKKLHTQERAVEEVK-----	1083
Qy	312	ALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKI-----	352
		: : : :	
Db	1084	-----LSIKPYYQKKDITKEEY-----KDILRKAVHKICHSKSGEINPVKVSNLVRA	1130
Qy	353	-----PGL	355
Db	1131	YVQRYRYFRKHGRKPGDPPGPPRPPKEPGPPDKGGPGL	1168

hypothetical protein SPAC23A1.17 - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T38236
 R;Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, February 1998
 A;Reference number: Z21780
 A;Accession: T38236
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1611 <MUR>
 A;Cross-references: EMBL:AL021813; PIDN:CAA16991.1; GSPDB:GN00066;
 SPDB:SPAC23A1.17
 A;Experimental source: strain 972h-; cosmid c23A1
 C;Genetics:
 A;Gene: SPDB:SPAC23A1.17
 A;Map position: 1

Query Match 14.6%; Score 265.5; DB 2; Length 1611;
 Best Local Similarity 10.1%; Pred. No. 3.1;
 Matches 131; Conservative 59; Mismatches 106; Indels 1005; Gaps 30;

Qy	2	EDIDQSSLVS-----SSTDSPPRP-----	20
		::: ::: :::	
Db	299	EELSKSQRVAKDDDPFVVSNTANSDEPASSSKPAKPLTDLNRAFSQRLNLDPOKPGKSQG	358
Qy	21	-----PPAFKYQFVTEPED-----EEDEEEEEDEEEE	46
		::: ::: :::	
Db	359	EISEQEEDEYDDAESDEMHSPLYSTHEPESEPEDQDEPSEKDDENKDVVEEQEQEQEQEQEQI	418
Qy	47	DDEDLLEELEVLER-----	59
		::: :::	
Db	419	DPEEAKRIALRERMAKMSGGIGMHVFGPLGLAAPIGRKNTLRRTPAKSSEEAKSTTNDSS	478
Qy	60	-----KPAAGLS	66
Db	479	PPKDSSSTSTQPTEQSNAQQAPSPKEEERPLPSEPSQNQPAEYRDTPTDTPRNIMPLPGLM	538
Qy	67	AA-----	68
		:	
Db	539	SADQPIKVTEPSNDADKAIVAEGPNNEEETKGPVIPETQETSEQQVHKTPSPEKQKVLSP	598
Qy	69	-----AVPPAAAAPLL-----	79
		::: :::	
Db	599	PPIITNFDKETLASNEAHEAVPQKPSAPQVTRLMAPQDSSSVVTPSPTSLLDPAVRKV	658
Qy	80	-----DFSS-----DS	85
		:::	
Db	659	IDGIDPPKEAGAGATADVESAANSPIPPRTWHSPDFTSKSFEPIERKLPSRISEVTEDS	718
Qy	86	V-----PPAP-----	90
		:	
Db	719	IDEDKQNEVDPST SARALPPPGLRFGKVDTLASLAHDDLDDLPVPRIFSPPLPKTPSG	778
Qy	91	-----RG-----P	93
Db	779	EFGDNEFMFPKKSNNRVRGHQSRPSTGSQLRNVVPVSIVTSGGRPALPDEMASPSSSIGHP	838

QY	94	LEAAPPA-----APERQPSWERS-----	111
		: :	
Db	839	LPSPPPADFNLSLNVDFYEPHSYLESPAPEPQPSYEEESFNATVIHAPTPSTATFQGHPTI	898
QY	112	-----	111
Db	899	SNVATPPLKQDVTESKASPVADASATHQSSTGLTQEITQLGSNMRLPTKLTRPSNDGRKA	958
QY	112	----PAAPAPSLPPAAAV-----L	126
		:	
Db	959	SGRPAAP-PSIPPPLPVSNILSSPTSEPPKDHPPSAPLSKPVSTSPAAPLARVPPVPKL	1017
QY	127	PSK-----LPEDDEPP-----	137
Db	1018	SSKAPPVPLPSADAPPPIVPSTAPPVPIPTSTPPVPKSSSGAPSAPPPVPAPSSEIPSIP	1077
QY	138	-----	137
Db	1078	APSGAPPVPAPSGIPPVPKPSVAAPPVPKPSVAVPPVPAPSGAPPVPKPSVAAPPVPVPS	1137
QY	138	-----ARPPPPPPAGASPLAEP--AAPP-----	158
		: : :	
Db	1138	GAPPVPKPSVAAPPVPAPSGAPPVPKPSVAAPPVPAPSSGIPPVPKPAAGVPPVPPPSEA	1197
QY	159	-----STP-----	161
Db	1198	PPVPKPSVGVPVPPPPSTAPPVPTPSAGLPVPVPTAKAPPVPAPSSEAPSVSTPRSSVP	1257
QY	162	-----AAPKR-----RGS	169
Db	1258	SPHSNASPSPTSSSMASAPARTSVSRSKSKAERHETSTSSRKSSKSGEHHHHHNEGHAD	1317
QY	170	GSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVF-----SIVSVTAYI-----	211
		: : :	
Db	1318	SSSTRTSLAHQDSRK-----SLHRHLSRSSSRASKKPSIVSTTGPFNESFSAKPVEPC	1370
QY	212	-----ALALLSVTISFRIYKGV-----	228
		:: :: :: : ::	
Db	1371	ASEKWWLNSTAVPKSVVQMNDVLYMIKEGITGQDKKYKSVHILFPDYSQTVLTATFNPH	1430
QY	229	-----IQAIQKSDEGHPFRAYLESEVAISEELVQK-----YNSALGHVNSTIKELRRLF	278
		: : : : : : ::	
Db	1431	NQNITQLSQLQLAPPAQPSKARLDEEYACYGSTILKKARAYQGSVMVGDSA-----	1481
QY	279	LVDDLVDLSLKFAVLMWVFTYVGAL-----FNGLTLLILALISLFSIPVIY	323
		: ::	
Db	1482	-----FTFVNSVMSILAHNLEPINKQTFGG-----VIY	1509
QY	324	ERHQVQIDHYLGLANKSVK-----DAMAKIQAKIPGLK	356
		: : : : :	
Db	1510	K-----NVGNVTVOOIGEIRPGDIVTFDKAKFSGOK	1540

RESULT 10

cappuccino gene protein - fruit fly (*Drosophila melanogaster*)


```

Qy      284 VDS----LKFAVLMWVFTTYVGALFNG-----LTLL-- 309
      :||  | |::: : |:  ||  |||
Db      811 IESEDLKLVFSIILTILGNYM----NGGNRQRGQADGFNLDILGKLKDVKSSESHTTLLHF 866

Qy      310 ----- 309

Db      867 IVRTYIAQRRKEGVHPLEIRLPIPEPADVERAAQMDFEVQQQIFDLNKKFLGCKRTTAK 926

Qy      310 ILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKI 352
      :||  | | | | : ::: : |::|  ||:  :
Db      927 VLAA----SRPEIMEFPFKSKMEEFVEGADKS----MAKLHQSL 961

```

RESULT 11

S24407

formin isoform IV - mouse

C;Species: Mus musculus (house mouse)

C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C;Accession: S24407

R;Jackson-Grusby, L.; Kuo, A.; Leder, P.

Genes Dev. 6, 29-37, 1992

A;Title: A variant limb deformity transcript expressed in the embryonic mouse limb defines a novel formin.

A;Reference number: S24407; MUID:92112033; PMID:1339380

A;Accession: S24407

A;Molecule type: mRNA

A;Residues: 1-1206 <JAC>

A;Cross-references: EMBL:X62379; NID:g51552; PIDN:CAA44244.1; PID:g51553

```

Query Match          14.1%;  Score 257.5;  DB 2;  Length 1206;
Best Local Similarity 14.1%;  Pred. No. 2.8;
Matches   95;  Conservative   39;  Mismatches   95;  Indels   445;  Gaps   22;

```

```

Qy      13 STDSPPRPPP-----AFKYQFV----- 29
      |||  |  |||  |:
Db      373 STDQESHKSPRDAHVQGGQVKARTPETALEAFKALFIRPPKKGSTADTSELEALKRKMKH 432

Qy      30 -----TEPEDEEDEEEEEDE---EEDD----- 48
      : | |  : : |:  |:||
Db      433 EKESLRAVFERSKSRPADSPSPDKSPDQSPTEQDDRTPGRLQAVWPPPKTKDTEEKVGLK 492

Qy      49 -----EDLEELEVLER----- 59
      |  ||:| |:
Db      493 YTEAEYQAAILHLKREHKEEIELQAQFELKTFHIRGEHALVTARLEEAIENLKQQLEKR 552

Qy      60 -----KPAAGLSAAA----- 69
      ||  | |
Db      553 REGCEEMRDVCISTDDDCSPKAFRNVCIQTDRETFLKPCDAESKATRSSQIVPKKLTISL 612

Qy      70 -----VPPAAAAPLLDFFSSDSVPPAP--- 90
      ||  ||  :|| |
Db      613 TQLSPSKDSKDIHAPFQTREGTSSSSQQKISPPAPPTPPPLPPPL-----IPPPPPPLP 665

Qy      91 --RGPLPAAPPAAPERQPSWERSPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPP--- 145
      |||| ||| |  | :| | ||  | : :| | ||  |||||
Db      666 PGLGPLPPAPPPIPV-----CPVSPPPPPPPPP---PTPVPPSDGPPPPPPPPPPPLPN 715

```

```

Qy      146 -----AGASP-----LAEPAAP-----PSTPAAPKRRG 168
          :| |           || | |           || | |
Db      716 VLALPNSGGPPPPPPPPPPGLAPPPPPGLSFGLSSSSSQYPRKPAIEPSCPMKP---- 771

Qy      169 SGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 228
          |||
Db      772 -----LYWTR----- 776

Qy      229 IQAIQKSDEGHP-----FRAYLESEVAISEELVQ----- 257
          ||  || : |           | | | | : : |
Db      777 IQINDKSQDAAPTLWDSLEEPHIRTSEFEYLFSDKDTTQQKKKPLSEAYEKKNKVKKIIK 836

Qy      258 ----KYSNSA---LGHVNSTIKELRR-LFLVDDLVDLSLKFAVLMWVFTYVGALENGLTLL 309
          | | :   : : : :| : : : | |||
Db      837 LLDGKRSQTVGILISSLHLEMKDIQQAIFTVDD----- 869

Qy      310 ILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQ----- 349
          : : : | : : : ||           | : : | : || :
Db      870 --SVVDLETLAALYE-----NRAQEDELTKIRKYETSKKEEDLKLLDKPEQF 914

Qy      350 ----AKIPGLKRKA 359
          | : ||   : |
Db      915 LHELAQIPNFAERA 928

```

RESULT 12

A59266

unconventional myosin-15 - human

C;Species: Homo sapiens (man)

C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 08-Sep-2000

C;Accession: A59266

R;Liang, Y.; Wang, A.; Belyantseva, I.A.; Anderson, D.W.; Probst, F.J.; Barber, T.D.; Miller, W.; Touchman, J.W.; Jin, L.; Sullivan, S.L.; Sellers, J.R.; Camper, S.A.; Lloyd, R.V.; Kachar, B.; Friedman, T.B.; Fridell, R.A. Genomics 61, 243-258, 1999

A;Title: Characterization of the human and mouse unconventional myosin XV genes responsible for hereditary deafness DFNB3 and shaker 2.

A;Reference number: A59266; MUID:20021762; PMID:10552926

A;Accession: A59266

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-3530 <LIA>

A;Cross-references: GB:AF144094; NID:g6224682; PIDN:AAF05903.1; PID:g6224683

C;Superfamily: myosin motor domain homology

F;1225-1887/Domain: myosin motor domain homology <MMO>

```

Query Match          14.1%; Score 256.4; DB 2; Length 3530;
Best Local Similarity 10.1%; Pred. No. 45;
Matches 145; Conservative 54; Mismatches 124; Indels 1106; Gaps 37;

```

```

Qy      17 PPRPP-----PAFKYQFV-----TEPEDEEDEEEE----- 40
          || |           || | :|           |||| | :
Db      394 PPEVPYFYPEESASAFVYPWPVPPPIPSPHNPYAHAMDDIAELEEPEDAGVERQGTSFRLP 453

Qy      41 -----EEDEE 45
          : |||
Db      454 SAAFFEQQGMDKPARSKLSLIRKFRLLFPRPVKLFQKEKLEVPLPPSLDIPLPLGDADEE 513

```

Qy	46	EDDEDLEELEVL-----	57
		: : : :	
Db	514	EDEEELPPVSAVPYGHFPWGFLLTPRQRNLQRALSAFGAHRGLGFGPEFGRPVPRPATSLA	573
Qy	58	-----ERKPAAGL-----	65
		:	
Db	574	RFLKKTLSEKKPIARLRGSQKTRAGGPAVREAAAYKRFQYKLAGMDPEKPGTPIVLRRAQP	633
Qy	66	-----	65
Db	634	RARSSNDARRPPAPQPAPRTLSHWSALLSPPVPPRPPSSGPPAPPLSPALSGLPRPASP	693
Qy	66	-----SAAAVPPAAAA-----	76
		:	
Db	694	YGSLRRHPPPWAAPAHVPPAPQASGWAFVEPPAVSPEVPPDLLAFPGPRPSFRGSRRRGA	753
Qy	77	-----PLLDFSSDSVPPAPR-----GPL--PAAPP	99
		: : :	
Db	754	AFGFPGASPRASRRRAWSPLASPQPSLRSSPGLGYCSPLAPPSPQLSLRTGPFQPPFLPP	813
Qy	100	A-----APERQ-----PSWERSPAAPAPSLPP-----AAAVLPSKL-----	130
		: : : :	
Db	814	ARRPRSLQESPAPRRAAGRLGPPGSPPLPGSPRPPSPPLGLCHSPRRSSINLPSRLPHTWR	873
Qy	131	-----PEDDE-----	135
Db	874	RLSEPPTRAVKPQVRLPFHRPPRAGAWRAPLEHRESPREPEDSETPWTVPPLAPSWDVDM	933
Qy	136	-PPARPPPPPPAGA-----SPLAEPAP-----	157
		:	
Db	934	PPTQRPPSPWPGGAGSRRGFSRPPVPENPFLQLLGPVPSPTLQPEDPAADMTRVFLGRH	993
Qy	158	-----PSTPAAPK-----	165
Db	994	HEPGPGQLTKSAGPTPEKPEEEATLGDPQLPAETKPPTPAPPKDVTTPPKDITPPKDVLP	1053
Qy	166	-----RR-----	167
Db	1054	QKTLRPSLSYPLAACDQTRATWPPWHRWGTLQAAAPLAPIRAPEPLPKGGERRQAAPGR	1113
Qy	168	-----	167
Db	1114	FAVVMPRVQKLSSFQRVGPATLKPQVQPIQDPKPRACSLRWSCWLRLADAYGPWPRVHTH	1173
Qy	168	-----GSGSS-----VVDLLYWRDI	182
		:: : :	
Db	1174	PQSCHLGPGAACLSLRGSWEEVGPPSWRNKMHSIRNLPSMRFREQHGEDGVEDMTQLEDL	1233
Qy	183	KKTGVVFGASLFLLLSLTVFSIVSVT-----AYIALALLSVT--ISFRIY-----	225
		: : : : :	
Db	1234	QET-----TVLSNLKIRFERNLIYTYIGSILVSVNPNPYQMFQIYGPEQVQY	1279
Qy	226	-----K	226
		:	
Db	1280	NGRALGENPPHLFAVANLAFKMLDAKQNCIIISGESGSGKTEATKLILRYLAAMNQKR	1339

Qy 227 GVIQAIQ-----KSDEGHPFRAYLE-----SEVAISEELVQK--- 258
 |:| |: :|| | :|| | |:| :||
 Db 1340 EVMQQIKILEATPLLESFGNAKTVRNDNSSRFGKFVEIFLEGGVISGAITSQYLLEKSRI 1399
 Qy 259 -----YSNSA----- 263
 | |
 Db 1400 VFQAKNERNYHIFYELLAGLPAQLRQAFSLQEAETYYYLNQGGNCEIAGKSDADDFRRL 1459
 Qy 264 -----LGH----- 266
 ||:
 Db 1460 AAMEVLGFSSSEDQDSIFRILASILHLGNVYFEKYETDAQEVASVVSAREIQAVAELLQIS 1519
 Qy 267 -----VNSTIKELRRLE-----LVD--DLVDSLKFAVLM-WVFTTYVGAL--- 302
 | |::| ::| || | : : :|:| |:| | |
 Db 1520 PEGQLKAITFKVTETMRE--KIFTPLTVESAVDARDAIAKVLVYALLFSWLITRVNALVSP 1577
 Qy 303 -----FNGLTLLIL-----ALISLFSIPVI-----YERHQVQ--- 329
 || | : | ||: | | |:
 Db 1578 RQDTLSIAILDIYGFEDLSFNSFEQLCINYANENLQYLENKIVFQEEQEEYIREQIDWQE 1637
 Qy 330 ----- 329
 Db 1638 ITFADNQPRINLISLKPYGILRILDDQCCFPQATDHTFLQKCHYHHGANPLYSKPKMPLP 1697
 Qy 330 ---IDHYLGLA----- 337
 | | |
 Db 1698 EFTIKHYAGKVITYQVHKFLDKNHDQVRQDVLDFVRSRTRVVAHLFSSHAPQAAPQRLGK 1757
 Qy 338 -----NKSVDAMAKIQ-----AKIPGL 355
 :|: | : |:| | |||
 Db 1758 SSSVTRLYKAHTVAAKFQQSLLDLVEKMERCNPLFMRLKPNHKKEPGL 1806

RESULT 13

S11515

formin - mouse

C;Species: Mus musculus (house mouse)

C;Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C;Accession: S11515

R;Woychik, R.P.; Maas, R.L.; Zeller, R.; Vogt, T.F.; Leder, P.

Nature 346, 850-853, 1990

A;Title: 'Formins': proteins deduced from the alternative transcripts of the limb deformity gene.

A;Reference number: S11515; MUID:90363291; PMID:2392150

A;Accession: S11515

A;Molecule type: mRNA

A;Residues: 1-1468 <WOY>

A;Cross-references: EMBL:X53599; NID:g52877; PIDN:CAA37668.1; PID:g52878

Query Match 14.0%; Score 254.9; DB 2; Length 1468;
 Best Local Similarity 14.5%; Pred. No. 5.7;
 Matches 97; Conservative 40; Mismatches 103; Indels 431; Gaps 23;

Qy 2 EDIDQSSLVSSSTDSPPRP---PPA-----FKYQ----- 27
 |: : | :||: || || |
 Db 602 EEASEKGLGPEKITAPPQHQLPFGIASEGFPCDNFKEQTAKDLPNKDGGVWVPGYRAGPP 661

QY	28 --FVTEPEDEEEDD-----EDEEDD-----	48
	: : :	
Db	662 CPFLLEEKEKTSRSELYLDLNPDSPTQDDRTPGRLQAVWPPPKTKDTEEKVGLKYTE	721
QY	49 -----EDLEEEVLER-----	59
	: :	
Db	722 AEYQAAILHLKREHKEEIETLQAQFELKTFHIRGEHALVTARLEEAIENLKQQLEKRREG	781
QY	60 -----KPAAGLSAAA-----	69
Db	782 CEEMRDVCISTDDDCSPKAFRNVCIQTDRETFLKPCDAESKATRSSQIVPKKLITISLTQL	841
QY	70 -----VPPAAAPLLDFSSDSVPAP----R	91
	:	
Db	842 SPSKDSDKIHAFQTREGTSSSSQQKISPPAPPPTPPLPPLL-----IPPPPLPPGL	894
QY	92 GPLPAAPPAAPERQPSWERSPAAPAPSLPPAAVLPSKLPEDDEPPARPPIPPP-----	145
	: : :	
Db	895 GPLPPAPPPIPPV-----CPVSPPPPPPPP---PTPVPPSDGPPPPPPPPPLPNVLA	944
QY	146 ---AGASP-----LAEPAAP-----PSTPAAPKRRSGS	171
	:	
Db	945 LPNSGGPPPPPPPPPPGLAPPPPPGLSFGLSSSSSQYPRKPAIEPSCPMKP-----	997
QY	172 SVVDLLYWWDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQA	231
Db	998 ----LYWTR-----IQUI	1005
QY	232 IQKSDEGHP-----FRAYLESEVAISEELVQ-----	257
	: : :	
Db	1006 NDKSQDAAPTLDWSLEEPHIRTSEFEYLFSKDTTQKKKPLSEAYEKKNVKKIILDL	1065
QY	258 -KYSNSA---LGHVNSTIKELRR-LFLVDDLVDLSKFVLMWFTYVGALFNGLTLILA	312
	: : : : : :	
Db	1066 GKRSQTVGILISSLHEMKDIQQAIFTVDD-----S	1096
QY	313 LISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQ-----	349
	: : : : : : : : :	
Db	1097 VVDLETALAALYE-----NRAQEDELTKIRKYETSKEEDLKLLDKPEQFLHE	1143
QY	350 -AKIPGLKRKA	359
	: :	
Db	1144 LAOIPNEAERA	1154

A;Residues: 1-716 <WIL>
A;Cross-references: EMBL:AL110490; NID:e1542263; PIDN:CAB54442.1; CESP:Y48B6A.6
A;Experimental source: clone Y48B6A
C;Genetics:
A;Gene: CESP:Y48B6A.6
A;Introns: 38/3; 196/3; 437/2; 460/3; 518/1; 617/3; 673/3

Query Match 14.0%; Score 254.7; DB 2; Length 716;
Best Local Similarity 16.9%; Pred. No. 0.97;
Matches 101; Conservative 39; Mismatches 106; Indels 353; Gaps 23;

```

Qy      5 DQSSILVSSSTD-----SPPRPP-----PAFKYQFVTEPEDEEDEEEDEEEDD 48
      : ||| | ||| || | : | | :|| | | :
Db     178 EMESLVQKQVDVLDQIMSSPPPPPTSPPLPPTPPARITSVRSDIESIEEEERRRKESEET 237

Qy     49 EDLEELEV----LERKPAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGPLP----- 95
      ||| : | | || | : | | | | |
Db     238 ASFEELEAEIMRISRSP-----VPP---PVL---SIPPPPPNIPPLPTIPQEVQSP 282

Qy     96 -----AAPPAAPERQPSWE-----RSPAAPAPSLPP----- 121
      : || | || : : : | | |
Db     283 PSRPRTSVPPPIPSPGPSEDVNMDELIESFSDSVIFNNSMSPPPLPPLRESSLETLEVT 342

Qy    122 -----AAAVLPSKLPE-----DDEPPA----- 138
      | : || : || | |
Db     343 PEDPVTESKVEASPTPLPKATESLNESSIKALEGLEVKALEAQEASDDRPSAPTPIRDSS 402

Qy    139 RPPPPPPAGASPLA-----EPAAP-----PSTPAAPKRR 167
      ||||| :|| | :|| | :|| |
Db     403 LPPPPPPPKPETPLAIRRAGPIPTPQLLEMIHQEDCSIRPSSPTSVSHGSRPQSPAVPKKP 462

Qy    168 GSGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS----- 221
      ||| : |
Db     463 -----SVTVSPLGLLC 473

Qy    222 -----FRIYKGVIAI- 232
      :| : :
Db     474 DPNLSIEKPEEMKTEDTKPVETAPAPVDEAELNDALDRRNKINEATCLTKIFPSLVSKYF 533

Qy    233 -----QKSDEGHP-----FRAYLESE----- 248
      :| | | | : | | :
Db     534 QNCSFDLNDKRSGENVPLKMNKKISLYAEFSEFSRKQIQYFSGIFKKYDEDQDSYIDFN 593

Qy    249 -----VAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVIM 293
      :| : || : | : | | : | | :
Db     594 ELKRMMEKLGEAQTHIAL-KELIKKVEDQDGKISQ-----REFFLIIFRLAASGE----- 642

Qy    294 WVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKI 352
      | : :| : | | : || | | |||
Db     643 -----LSCSEVFK--TLAESVDVSKEGVLGAAN-----FFQAKI 674

```

RESULT 15

F86387

probable Pto kinase interactor [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F86387

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-760 <STO>

A;Cross-references: GB:AE005172; NID:q11079512; PIDN:AAG29223.1; GSPDB:GN00141

C; Genetics:

```
A;Map position: 1
```

Query Match 13.9%; Score 253.8; DB 2; Length 760;
Best Local Similarity 17.1%; Pred. No. 1.2;
Matches 115; Conservative 42; Mismatches 132; Indels 382; Gaps 25;

QY	2	EDIDQS-SLVSSSTDSPRRPPPAFKYQFVTEPEDEEEDDEEEEDDEDLEELEVLERK	60
		:: : : :	
Db	10	EEVSLSPSLASPPMLALPPPQPSFPGDNATSP-----TRE	44
QY	61	PAAGLSAAAV-PPAAAAPLLLDFSSDSVPPAPRGPLP-----	95
		:: :	
Db	45	PTNGNPPETTNTPAQSSPPPETPLSSPPPEPSPSPSLTGPPPTTIPVSPPPEPSPPPPL	104
QY	96	---AAPPAAPERQPSWERSPAAP-----APSLP-----	120
Db	105	PTEAPPANPVSSPPPESSPPPPPPTEAPPTTPTITSPSPPTNPPPPESPPSLPAPDPPS	164
QY	121	-----PAAAVLPSKLPEDDEPPARPP---	141
		: :	
Db	165	NPLPPPCLVPSSHSPRHLPSPPASEI PPPPRHLPSPPASERPSTPPSDSEHPSPPPHGH	224
QY	142	-----PPPPAGASPLAEPAAP-----PSTPAAPKR-----	166
		: : :	
Db	225	PKRREQPPPPGSKRTPSPSPSPDSKRPVHPSPSPPEETLPPPKPSPDPLPSNSSSPPT	284
QY	167	-----RGS GSSVVDLLYWDRDIKKTGVVFG	190
		: ::	
Db	285	LLPPSSVVSPSPPRKSVSGPDNPSPNNPTPVTDNSSSSGISIAAVV-----GVSIG	336
QY	191	ASLFLLLSLTVFIV-----	205

```

      :| |   ||:  :|
Db    337 VALVL---LTLIGVVVCCCLKRKKRLSTIGGGYVMPTPMESSSPRSDSALLKTQSSAPLV 393

Qy    206 ----SVTAYIALA-----LLSVTISF-----RIYKGV----- 228
      |   |::  :           |:  |   |           |:| | |
Db    394 GNRSSNRTYLSQSEPGGFGQSRELFSEELVIATNGFSDENLLGEGGFGRVYKGVLPDER 453

Qy    229 IQAIQKSDEG-----HPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVD-- 281
      : |:::  |           |:| :::           : : :  : |   | | | | : |
Db    454 VVAVKQLKIGGGQGDREFKAEVDT-----ISRVHHRNLLSMVGYCISENRRLLIYDYV 506

Qy    282 -----DLVDSLKFAVLMWVFYVGALENGLTLL----- 309
      |           :| |           ||   ||   |
Db    507 PNNNLYFHLHGTPGLDWATRVKIA-----AGAA-RGLAYLHEDCHPRIIHRDIKSSN 557

Qy    310 IL-----ALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPG----- 354
      ||           ||:| |           ||| |   |   | : : |
Db    558 ILLENNFHALVSDF-----GLA-KLALDCNTHITTRVMGTFGYMAPEY 599

Qy    355 -----LKRKAD 360
      |   |:|
Db    600 ASSGKLTEKSD 610

```

Search completed: September 29, 2004, 18:55:42
Job time : 35.4632 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:40:14 ; Search time 76.3939 Seconds
(without alignments)
1486.854 Million cell updates/sec

Title: SEQ2_1-171_975-1163
Perfect score: 1823
Sequence: 1 MEDIDQSSLVSSSTDSPRP.....VKDAMAKIQAKIPGLKRKAD 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query				
No.	Score	Match	Length	DB	ID
Description					

1	1721.6	94.4	356	11	Q8BH78	Q8bh78 mus musculu
2	1712.7	93.9	375	11	Q8BHF5	Q8bhf5 mus musculu
3	1711.5	93.9	357	11	Q8K3G7	Q8k3g7 mus musculu
4	1637	89.8	1162	11	Q8BGM9	Q8bgm9 mus musculu
5	1622.9	89.0	1163	11	Q8K3G8	Q8k3g8 mus musculu
6	1598.4	87.7	392	4	Q96B16	Q96b16 homo sapien
7	1012.1	55.5	986	4	Q8IUA4	Q8iua4 homo sapien
8	1010.2	55.4	1046	11	Q8BGK7	Q8bgk7 mus musculu
9	992.2	54.4	578	11	Q8OW95	Q8ow95 mus musculu
10	988.7	54.2	639	11	Q8K290	Q8k290 mus musculu
11	878	48.2	184	6	Q7YRW9	Q7yrw9 bos taurus
12	874.8	48.0	199	13	Q7T224	Q7t224 gallus gall
13	855.6	46.9	720	11	Q7TNB7	Q7tnb7 mus musculu
14	844	46.3	179	6	Q9GM33	Q9gm33 macaca fasc
15	826.2	45.3	780	11	Q8K0T0	Q8k0t0 mus musculu
16	823.2	45.2	780	11	Q8K4S4	Q8k4s4 mus musculu
17	793.9	43.5	760	13	Q90638	Q90638 gallus gall
18	736.6	40.4	214	13	Q7T222	Q7t222 carassius a
19	687.7	37.7	643	11	Q8CCU2	Q8ccu2 mus musculu
20	683	37.5	199	4	Q9BQ59	Q9bq59 homo sapien
21	675.4	37.0	595	5	Q9VMV9	Q9vmv9 drosophila
22	669	36.7	267	11	Q63765	Q63765 rattus sp.
23	665	36.5	208	13	Q90637	Q90637 gallus gall
24	659.9	36.2	237	11	Q8C6D5	Q8c6d5 mus musculu
25	657.4	36.1	236	11	Q8VBU0	Q8vbu0 rattus norv
26	588.6	32.3	221	13	Q7ZUD6	Q7zud6 brachydanio
27	527.7	28.9	222	5	Q9VMW4	Q9vmw4 drosophila
28	523.9	28.7	234	5	Q9VMW3	Q9vmw3 drosophila
29	521	28.6	224	5	Q9VMW1	Q9vmw1 drosophila
30	518	28.4	202	5	Q9VMW2	Q9vmw2 drosophila
31	437.7	24.0	2484	5	Q9U347	Q9u347 caenorhabdi
32	427.4	23.4	2607	5	Q23187	Q23187 caenorhabdi
33	344	18.9	107	13	Q7T223	Q7t223 carassius a
34	337.5	18.5	222	5	Q23188	Q23188 caenorhabdi
35	286.9	15.7	1312	4	Q7Z5V7	Q7z5v7 homo sapien
36	286.9	15.7	1343	4	Q9H7N4	Q9h7n4 homo sapien
37	278.9	15.3	1312	4	Q9NR59	Q9nr59 homo sapien
38	272.4	14.9	1217	4	Q9ULL5	Q9ull5 homo sapien
39	266.9	14.6	1173	11	Q63624	Q63624 rattus norv
40	265.5	14.6	1611	3	O42854	O42854 schizosacch
41	263.6	14.5	1179	12	Q91L98	Q91l98 white spot
42	263.5	14.5	1180	12	Q8VAS9	Q8vas9 white spot
43	263.4	14.4	1709	4	O15047	O15047 homo sapien
44	263	14.4	1185	12	Q8QTC5	Q8qtc5 white spot
45	263	14.4	1627	10	Q84ZL0	Q84zl0 oryza sativ

ALIGNMENTS

RESULT 1

Q8BH78

ID Q8BH78 PRELIMINARY; PRT; 356 AA.
AC Q8BH78;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE RTN4.
 GN RTN4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvcJ7, and 129SvcJ7;
 RA Oertle T., van der Putten H., Schwab M.E.;
 RT "Genomic Structure and Functional Characterization of the Promoter
 RT Structures of Human and Mouse Nogo/Rtn-4.";
 RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvcJ7, and 129SvcJ7;
 RA Oertle T., Schwab M.E.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvcJ7;
 RA Van der Putten H.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129SvcJ7;
 RA Van der Putten H., Mir A.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AY102281; AAM73503.1; -.
 DR EMBL; AY102286; AAM73508.1; -.
 DR MGD; MGI:1915835; Rtn4.
 DR GO; GO:0005783; C:endoplasmic reticulum; IDA.
 DR GO; GO:0007399; P:neurogenesis; IDA.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 SQ SEQUENCE 356 AA; 38403 MW; 4366C03BA9630B56 CRC64;

Query Match 94.4%; Score 1721.6; DB 11; Length 356;
 Best Local Similarity 96.4%; Pred. No. 1.3e-63;
 Matches 347; Conservative 2; Mismatches 7; Indels 4; Gaps 3;

Qy	1	MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK	60
Db	1	MEDIDQSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEEDEE-EEDEDEDLEEEVLERK	59
Qy	61	PAAGLSAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	60	PAAGLSAAPVPP-AAAPLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP	118
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPSTPAAPKRRGSGSSVVDLLYWR	180
Db	119	PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAPSTPAAPKRRGSGSVVVDLLYWR	176
Qy	181	DIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHP	240
Db	177	DIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHP	236

QY 241 FRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLQKFAVLMWVFTYVG 300
 |||
 Db 237 FRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLQKFAVLMWVFTYVG 296
 QY 301 ALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
 |||
 Db 297 ALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAE 356

RESULT 2

Q8BHF5

ID Q8BHF5 PRELIMINARY; PRT; 375 AA.
 AC Q8BHF5;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE RTN4.
 GN RTN4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvcJ7, and 129SvcJ7;
 RA Oertle T., van der Putten H., Schwab M.E.;
 RT "Genomic Structure and Functional Characterization of the Promoter
 RT Structures of Human and Mouse Nogo/Rtn-4.";
 RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvcJ7, and 129SvcJ7;
 RA Oertle T., Schwab M.E.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvcJ7;
 RA Van der Putten H.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129SvcJ7;
 RA Van der Putten H., Mir A.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AY102282; AAM73504.1; -.
 DR EMBL; AY102286; AAM73509.1; -.
 DR MGD; MGI:1915835; Rtn4.
 DR GO; GO:0005783; C:endoplasmic reticulum; IDA.
 DR GO; GO:0007399; P:neurogenesis; IDA.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 SQ SEQUENCE 375 AA; 40300 MW; 23D9EB19BE671AE6 CRC64;

Query Match 93.9%; Score 1712.7; DB 11; Length 375;
 Best Local Similarity 91.6%; Pred. No. 3.5e-63;

SO SEQUENCE 357 AA; 38566 MW; 73BB3D17DFDBDF15 CRC64;

Qy	1	MEDIDQSSLVSSSTDSPPRPPAFKYQFVTEPEDEEDEEEDEEDEDLEEEVLERK	60
Db	1	MEDIDQSSLVSSSADSPPRPPAFKYQFVTEPEDEEDEEEDEE-EEEDDEDLEEEVLERK	59
Qy	61	PAAGLSAAAVPPAAAAPLLDfSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	60	PAAGLSAAPVPP-AAAPLLDfSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP	118
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPSTPAAPKRRGSGSSVVDLLYWR	180
Db	119	PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAAPSTPAAPKRRGSGSVVDLLYWR	176
Qy	181	DIKKTGVV-FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGH	239
Db	177	DIKKTGVVYFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGH	236
Qy	240	PFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYV	299
Db	237	PFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYV	296
Qy	300	GALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKA	359
Db	297	GALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKA	356
Qy	360	D 360	
	:		
Db	357	E 357	

RESULT 4

```

ID      Q8BGM9          PRELIMINARY;          PRT;   1162 AA.
AC      Q8BGM9;
DT      01-MAR-2003   (TrEMBLrel. 23, Created)
DT      01-MAR-2003   (TrEMBLrel. 23, Last sequence update)
DT      01-OCT-2003   (TrEMBLrel. 25, Last annotation update)
DE      RTN4.
GN      RTN4.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=129/SvcJ7, and 129SvcJ7;
RA      Oertle T., van der Putten H., Schwab M.E.;
RT      "Genomic Structure and Functional Characterization of the Promoter
RT      Structures of Human and Mouse Nogo/Rtn-4.";
RL      Submitted (OCT-2002) to the EMBL/GenBank/DBDJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.

```


Db	417	DSESRNENASFPRTPELVKDGSRAYITCDSFSSATESTAANIFPVLEDHTSENKTDEKKI	476
Qy	172	-----	171
Db	477	EERKAQIITEKTSPKTSNPFLVAIHDSEADYVTTDNLSKVTEAVVATMPEGLTPDLVQEA	536
Qy	172	-----	171
Db	537	CESELNEATGTKIAYETKVDLVQTSEAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEA	596
Qy	172	-----	171
Db	597	PLNSLLPSTGASVAQPSASPLEVPSVSYDGIKLEPENPPPYEEAMSVALKTSDSKEEIK	656
Qy	172	-----	171
Db	657	EPESFNAAAQEAEPYISIIACDLIKETKLSTEPSPEFSNYSEIAKFEKSVPDHCELVDDS	716
Qy	172	-----	171
Db	717	SPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVSETVTQHKHKERLSASPQEVGKPY	776
Qy	172	-----	171
Db	777	LESFQPNLHITKDAASNEIPTLTKKETISLQMEEFNTAIYSNDDLSSKEDKMKESETFS	836
Qy	172	-----	171
Db	837	DSSPIEIIDEFPTFVSAKDDSPKEYTDLEVSNKSEIANVQSGANSLPCSELPCDLSFKNT	896
Qy	172	-----	171
Db	897	YPKDEAHVSDEFKSRSSSVSKVPLLLPNVSALESQIEMGNIVKPKVLTKEAEKLPDTE	956
Qy	172	-----SVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALA	214
Db	957	KEDRSLTAVLSAELNKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALA	1016
Qy	215	LLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKEL	274
Db	1017	LLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKEL	1076
Qy	275	RRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYL	334
Db	1077	RRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYL	1136
Qy	335	GLANKSVKDAMAKIQAKIPGLKRKAD	360
		:	
Db	1137	GLANKSVKDAMAKIQAKIPGLKRKAE	1162

RESULT 5

Q8K3G8

ID Q8K3G8 PRELIMINARY; PRT; 1163 AA.
 AC Q8K3G8;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Nogo-A.
 GN RTN4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Jin W., Long M., Li R., Ju G.;
 RT "Cloning and expression of the mouse Nogo-A protein.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AY114152; AAM77068.1; -.
 DR MGD; MGI:1915835; Rtn4.
 DR GO; GO:0005783; C:endoplasmic reticulum; IDA.
 DR GO; GO:0007399; P:neurogenesis; IDA.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 SQ SEQUENCE 1163 AA; 126691 MW; 6B5F362799417EA4 CRC64;

Query Match 89.0%; Score 1622.9; DB 11; Length 1163;
 Best Local Similarity 29.7%; Pred. No. 3.7e-58;
 Matches 347; Conservative 2; Mismatches 7; Indels 811; Gaps 5;

Qy	1	MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEELEVLERK	60
Db	1	MEDIDQSSLVSSSADSPRRPPPAFKYQFVTEPEDEEDEEDEE-EEEDDEDLEELEVLERK	59
Qy	61	PAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	60	PAAGLSAVPVPP-AAAPLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP	118
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPSTPAAPKRRGSGS-----	171
Db	119	PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAPSTPAAPKRRGSGSVDETLFALP	176
Qy	172	-----	171
Db	177	AASEPVIPISSAEKIMDLKEQPGNTVSSGQEDFPSVLFETAASLPSLSPLSTVSFKEHGYL	236
Qy	172	-----	171
Db	237	GNLSAVASTEGTIEETLNEASRELPERATNPFVNRESAEFSVLEYSEMGSFNGSPKGES	296
Qy	172	-----	171
Db	297	AMLVENTKEEVIVRSKDKEDLVCSAALHNPQESPATLTKVVKEDGVMSPEKTMDFNEMK	356
Qy	172	-----	171
Db	357	MSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANMESKVDKKCFEDSLEQKSHGK	416
Qy	172	-----	171
Db	417	DSESRNENASFPSTPELVKDGSRAYITCDSFTSATESTANIFPVLEDHTSENKTDEKKI	476

Qy	172	-----	171
Db	477	EERKAQIITEKTS PKTSNPFLVAIH DSEADYVTTDNLSKVTEAVVATMPEGLTPDLVQEA	536
Qy	172	-----	171
Db	537	CESELNEATGTKIAYETKVDLVQTSEAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEA	596
Qy	172	-----	171
Db	597	PLNSLLPSTGASVAQPSASPLEVSPVSYDGIKLEPENPPPYEEAMSVALKTS DAK EEIK	656
Qy	172	-----	171
Db	657	EPESFNAAAQAEAPYIS IACDLIKETKLSTEPSPGFSNYSEIAKFEKSVPDHC ELVDDS	716
Qy	172	-----	171
Db	717	SPESEPVDLFSDDSIPEVPQTQEEAVMLMKESL TEVSETVTQHKHKERLSASPQEVGKPY	776
Qy	172	-----	171
Db	777	LESFQPNLHITKDAASNEIPTLT KKETISLQMEEFNTAIYSNDDLSSKEDKMKESETFS	836
Qy	172	-----	171
Db	837	DSSPIEIIDEFPTFVSAKDDSPKEYTDLEVSNKSEIANVQSGANSLPCSELPCDLSFKNT	896
Qy	172	-----	171
Db	897	YPKDEAHVSDEFKSKRSSVSKVPLLLPNVSALESQIEMGNIVKPKVLTKEAE EKLP SDTE	956
Qy	172	-----SVVDLLYWRDIKKTGVV-FGASLFLLLSLT VFSIVSVTAYIAL	213
Db	957	KEDRSLTAVLSAELNKT SVVDLLYWRDIKKTGVVYFGASLFLLLSLT VFSIVSVTAYIAL	1016
Qy	214	ALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKE	273
Db	1017	ALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKE	1076
Qy	274	LRRLFLVDDLVD SLKFVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHY	333
Db	1077	LRRLFLVDDLVD SLKFVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHY	1136
Qy	334	LGLANKSVKDAMAKIQAKIPGLKRKAD	360
Db	1137	LGLANKSVKDAMAKIQAKIPGLKRKAE	1163

RESULT 6

Q96B16

ID	Q96B16	PRELIMINARY;	PRT;	392 AA.
AC	Q96B16;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			

QY 167 RGS-----GSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVS 206
 ||| | : |||||
 Db 179 RGSSGSVDETLFALPAASEPVIRSSAVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVS 238
 QY 207 VTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGH 266
 |||||
 Db 239 VTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGH 298
 QY 267 VNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERH 326
 || ||||| : |||||
 Db 299 VNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERH 358
 QY 327 QVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
 | ||||| : |||||
 Db 359 QAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 392

RESULT 7

Q8IUAA4

ID Q8IUAA4 PRELIMINARY; PRT; 986 AA.
 AC Q8IUAA4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE RNT4 (RTN4).
 GN RTN4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oertle T., van der Putten H., Schwab M.E.;
 RT "Genomic Structure and Functional Characterization of the Promoter
 RT Structures of Human and Mouse Nogo/Rtn-4.";
 RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Oertle T., Schwab M.E.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Van der Putten H.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=22376540; PubMed=12488097;
 RA Oertle T., van der Putten H., Schwab M.E.;
 RT "Genomic Structure and Functional Characterization of the Promoter
 RT Structures of Human and Mouse Nogo/Rtn-4.";
 RL J. Mol. Biol. 325:299-323(2003).
 DR EMBL; AY102285; AAM64244.1; -.
 DR EMBL; AY123245; AAM64249.1; -.
 DR EMBL; AY123246; AAM64250.1; -.
 DR EMBL; AY123247; AAM64251.1; -.

DR EMBL; AY123248; AAM64252.1; -.
 DR EMBL; AY123249; AAM64253.1; -.
 DR EMBL; AY123250; AAM64254.1; -.
 DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 SQ SEQUENCE 986 AA; 108449 MW; 0CDE8F647036415A CRC64;

Query Match 55.5%; Score 1012.1; DB 4; Length 986;
 Best Local Similarity 27.7%; Pred. No. 7.3e-33;
 Matches 250; Conservative 33; Mismatches 71; Indels 549; Gaps 16;

Qy	3	DIDQSSLVSSSTDSPRPAPAFKYQFVTEPEDE-----EDEEE-----	40
		::: : : : :	
Db	88	ELEYSEMGSFSVS----PKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTA	143
Qy	41	-----EEDE-----EEDDED-----	50
		: : :	
Db	144	LTKLVKEDEVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAG	203
Qy	51	-----LEEL-----EVLERKPAAGLSAAVPP	72
		: : : : :	
Db	204	GKIESNLESKVDDKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNP	263
Qy	73	AA-----AAPLLDFSSDS----	85
		: :	
Db	264	AATESIATNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDY	323
Qy	86	-----VPPAPRGPLPAAPPAAPERQ-----	105
		:	
Db	324	VTTDNLTQVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSSEVMQES	383
Qy	106	-----PSWERSPAAPAPSLPP-----AAVLPSKLP-----	131
		: : : :	
Db	384	LYPAAQLCPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESI	443
Qy	132	--EDDEPP-----A	138
		:	
Db	444	KHEPENPPPYEEAMSVSLKKVSGIKKEEIKPENINAALQETEAPYISIIACDLIKETKLSA	503
Qy	139	RPPP-----PPPAGASPLAEPAPSTPA-----APKRR-----	167
		: : :	
Db	504	EPAPDFSDYSEMAKVEQVPVDHSELVEDSSPDSEPVDLFSDDSIQDVPQKQDETVMVLVKE	563
Qy	168	-----	167
Db	564	SLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIP	623
Qy	168	-----	167
Db	624	LQMEELSTAVYSNDDLFIKSEAQIRETETFSDDSPIEIIDEFPTLISSKTDSFSKLAREY	683
Qy	168	-----GSGS-----	171
		:	
Db	684	TDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSNGSATSQVL	743

Qy 172 -----SVVDLL 177
 Db 744 LLPPDVSA LATQAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLL 803
 Qy 178 YWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDE 237
 Db 804 YWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDE 863
 Qy 238 GHPPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFT 297
 Db 864 GHPPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFT 923
 Qy 298 YVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKR 357
 Db 924 YVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKR 983
 Qy 358 KAD 360
 Db 984 KAE 986

RESULT 8

Q8BGK7

ID Q8BGK7 PRELIMINARY; PRT; 1046 AA.
 AC Q8BGK7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE RTN4.
 GN RTN4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvcJ7, and 129SvcJ7;
 RA Oertle T., van der Putten H., Schwab M.E.;
 RT "Genomic Structure and Functional Characterization of the Promoter
 RT Structures of Human and Mouse Nogo/Rtn-4.";
 RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvcJ7, and 129SvcJ7;
 RA Oertle T., Schwab M.E.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvcJ7;
 RA Van der Putten H.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129SvcJ7;
 RA Van der Putten H., Mir A.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AY102280; AAM73502.1; -.

DR EMBL; AY102286; AAM73507.1; -.
DR MGD; MGI:1915835; Rtn4.
DR GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR GO; GO:0007399; P:neurogenesis; IDA.
DR InterPro; IPR003388; Reticulon.
DR Pfam; PF02453; Reticulon; 1.
DR PROSITE; PS50845; RETICULON; 1.
SQ SEQUENCE 1046 AA; 114221 MW; 8CE2E2238ED51222 CRC64;

Query Match 55.4%; Score 1010.2; DB 11; Length 1046;
Best Local Similarity 29.2%; Pred. No. 1e-32;
Matches 249; Conservative 25; Mismatches 80; Indels 498; Gaps 14;

```

Qy      5 DQSSLVSSSTDSPPRPPPA-----FKYQFVTEPEDE----- 35
      |: || |: |: || | | ||
Db    197 DKEDLVCSAALHNPQESPATLTKVVKEDGVMSPEKTMDFNEMKMSVVAPVREEYADFKP 256

Qy      36 -----EDEEEEEDEEDED----- 50
      || |: :| |
Db    257 FEQAWEVKDTYEGSRDVLAAANMESKVDKKCFEDSLEQKGHGKDSERNENASFPRTPE 316

Qy      51 -----LEELEVLERKPAAGLSAAVPP 72
      :| :: || | : |
Db    317 LVKDGSRAYITCDFSSATESTAANIFPVLEDHTSENKTDEKKIEERK--AQIITEKTSP 374

Qy      73 AAAAPLLDFSSDS-----VPPAPRGPLPAAPPAAPERQ----- 105
      :| | || | | | | | | |
Db    375 KTSNPFLVAIHDSEADYVTTDNLSKVTEAVVATMPEGLTPDLVQEACESELNEATGTKIA 434

Qy     106 -----PSWERSPAAPAPSLP-----PAAAVLPSKLPEDD 134
      ||:| :| |:| || | :|||
Db    435 YETKVDLVQTSEAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSLLPSTGASVA 494

Qy     135 EP-----PARPPP----- 142
      :| | ||
Db    495 QPSASPLEVPSPVSYDGIKLEPENPPPYEEAMSVALKTSDSKEEIKEPESFNAAAQEAEA 554

Qy     143 -----PPPA-----GASPLAEP-----A 156
      | | :|| :|| :
Db    555 PYISIACDLIKETKLSTEPSPEFSNYSEIAKFEKSVPDHCELVDSSPESEPVDFSDDS 614

Qy     157 PPSTP-----AAPKRRG----- 168
      | | |:| |
Db    615 IPEVPQTQEEAVMLMKESLTEVSETVTQHKHKERLSASPQEVGKPYLESFQPNLHITKDA 674

Qy     169 ----- 168

Db    675 ASNEIPTLTKKETISLQMEEFNTAIYSNDDLSSKEDKMKESETFSDSSPIEIIDFPTF 734

Qy     169 -----SG----- 170
      ||
Db    735 VSAKDDSPKEYTDLEVSNKSEIANVQSGANSLPCSELPCDLSFKNTYPKDEAHVSDEFK 794

Qy     171 ----- 170

Db    795 SRSSVSKVPLLLPNVSALESQIEMGNIVKPKVLTKEAEEKLPSDTEKEDRSLTAVLSAEL 854

```

Qy 171 --SSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 228
 :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 855 NKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 914

Qy 229 IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLK 288
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 915 IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLK 974

Qy 289 FAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKI 348
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 975 FAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKI 1034

Qy 349 QAKIPGLKRKAD 360
 |||||||||||:
 Db 1035 QAKIPGLKRKAE 1046

RESULT 9

Q80W95

ID Q80W95 PRELIMINARY; PRT; 578 AA.
 AC Q80W95;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Nogo-A (Fragment).
 GN NOGO-A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tozaki H., Hirata T.;
 RT "The partial sequence of mouse nogo-A cDNA clone#4109.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AB073672; BAC75974.1; -.
 DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 578 AA; 63696 MW; 832670C171E4AC61 CRC64;

Query Match 54.4%; Score 992.2; DB 11; Length 578;
 Best Local Similarity 42.4%; Pred. No. 1.2e-32;
 Matches 233; Conservative 13; Mismatches 79; Indels 225; Gaps 8;

Qy 17 PPRPPP---AFKYQFVTEPEDEEDEEEEEDEEEDDE-----DLEELEVLERKPAA 63
 | ||| | | ||:| | | : | :|:
 Db 48 PENPPPYEEAMSVALKTSDAKEEIKEPESFNAAQAQAEAPYISIACDLIKETKLSTEPSP 107

Qy 64 GLS-----AAAVP-----PAAAAPLLDFSSDSVPPAPR----- 91
 | | :|| :|: ||||:| |:
 Db 108 GFSNYSEIAKFEKSVDPDHCELVDSSPESEPVDFSDDSIPEVPQTQEEAVMLMKESLTE 167

Qy 92 -----GPLPAAPPAAPERQP----- 106
 | | ||

[illegible]

Query Match 54.2%; Score 988.7; DB 11; Length 639;
Best Local Similarity 42.2%; Pred. No. 2.2e-32;
Matches 232; Conservative 14; Mismatches 79; Indels 225; Gaps 8;

Qy	17	PPRPPP---AFKYQFVTEPEDEEDEEEEEDEEDDE-----DLEELEVLERKPAA	63
Db	109	PENPPPPYEEAMSVALKTSDAKEEIKEPESFNAAAQEAEPYISIACDLIKETKLSTEPS	168
Qy	64	GLS-----AAAVP-----PAAAPLLDFSSDSVPPAPR-----	91
Db	169	EFSNYSEIAKFEKSVDPDHCELVDSSPESEPVDFLSDDSIPEVPQTQEEAVMLMKESL	228
Qy	92	-----GPLPAAPPAAPERQP-----	106
Db	229	VSETVTQHKHKERLSASPQEVGKPYLESFQPNLHITKDAASNEIPTLTKKETISLQMEEF	288
Qy	107	-----	106
Db	289	NTAIYSNDDLSSKEDKMKESETFSDSSPIEIIIDFPTFVSAKDDSPKEYTDLEVS	348
Qy	107	-----SWERSPAAPAPSLPPAAAVLPS-	128
Db	349	IANVQSGANSLPCSELPCDLSFKNTYPKDEAHVSDEFKSRSSSVKVPLLLPNVSALESQ	408
Qy	129	-----KLPEDDEPPARPPPPPPAGASPLAEPAPPSTPAAPKRRGSG	170
Db	409	IEMGNIVKPKVLTKEAEKLPDTEKEDR-----SLTAVLSAELNK	449
Qy	171	SSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV	230
Db	450	TSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV	509
Qy	231	AIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVD	290
Db	510	AIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVD	569
Qy	291	VLMWVFETYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQA	350
Db	570	VLMWVFETYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLGLANKSVKDAMAKIQA	629
Qy	351	KIPGLKRKAD	360
Db	630	KIPGLKRKAE	639

RESULT 11

Q7YRW9

ID Q7YRW9 PRELIMINARY; PRT; 184 AA.

AC 07YRW9;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE RTN4w (Fragment).

GN RTN4.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22715887; PubMed=12832288;
 RA Oertle T., Klinger M., Stuermer C.A., Schwab M.E.;
 RT "A reticular rhapsody: phylogenic evolution and nomenclature of the
 RT RTN/Nogo gene family.";
 RL FASEB J. 17:1238-1247(2003).
 DR EMBL; AY164744; AAP47319.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 184 AA; 20671 MW; DE990E03BBAF84A1 CRC64;

Query Match 48.2%; Score 878; DB 6; Length 184;
 Best Local Similarity 96.2%; Pred. No. 3.6e-29;
 Matches 177; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 177 LYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSD 236
 |||
 Db 1 LYWRDIKKTGVVFGXXLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSD 60

 Qy 237 EGHPPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVF 296
 |||
 Db 61 EGHPPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVF 120

 Qy 297 TYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLK 356
 |||
 Db 121 TYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLK 180

 Qy 357 RKAD 360
 |||
 Db 181 RKAE 184

RESULT 12

Q7T224

ID Q7T224 PRELIMINARY; PRT; 199 AA.
 AC Q7T224;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE RTN4-C.
 GN RTN4.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22715887; PubMed=12832288;
 RA Oertle T., Klinger M., Stuermer C.A., Schwab M.E.;
 RT "A reticular rhapsody: phylogenic evolution and nomenclature of the
 RT RTN/Nogo gene family.";
 RL FASEB J. 17:1238-1247(2003).
 DR EMBL; AY164737; AAP47312.1; -.

SQ SEQUENCE 199 AA; 22293 MW; 07CF4E4EF2723251 CRC64;

Query Match 48.0%; Score 874.8; DB 13; Length 199;
Best Local Similarity 76.8%; Pred. No. 6.1e-29;
Matches 175; Conservative 11; Mismatches 12; Indels 30; Gaps 1;

```
Qy      133 DDEPPARPPPPPPAGASPLAEPAAPSTPAAPKRRGSGSSVVDLLYWRDIKKTGVVFGAS 192
          | :|                               | |||||
Db      2   DSQP-----SGWKDKVVDLLYWRDIKKTGVVFGAS 31

Qy      193 LFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAIS 252
          |||||
Db      32 LFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESDVAVS 91

Qy      253 EELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILA 312
          |::|||: |||:| :|||
Db      92 EDLIQKYSSVVLGHINGTVKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILA 151

Qy      313 LISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
          |||||:||||| ||||| ||:|||||
Db      152 LISLFSVPVIYERHQAQIDHYLGLVNKNVKDAMAKIQAKIPGLKRKTE 199
```

RESULT 13

Q7TNB7

ID Q7TNB7 PRELIMINARY; PRT; 720 AA.
AC Q7TNB7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BC056373; AAH56373.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 720 AA; 77435 MW; 80AB78728F16EAB2 CRC64;

Query Match 46.9%; Score 855.6; DB 11; Length 720;
 Best Local Similarity 46.9%; Pred. No. 1.1e-26;
 Matches 210; Conservative 28; Mismatches 66; Indels 144; Gaps 18;

```

Qy      1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEDEEDEDLEEEVLERK 60
          |||||
Db      1 MEDIDQSSLVSSSADSPRRPPPAFKYQFVTEPEDEEDEEDEE-EEDEDEDLEEEVLERK 59

Qy     61 PAAGLSAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120
          |||||
Db     60 PAAGLSAAPVPP-AAAPLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP 118

Qy    121 PAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPSTPAAPKRRGSGSSVVDLLYWR 180
          |||||
Db    119 PAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAPSTPAAPKRRGSGS--VD----- 169

Qy    181 DIKKTGVVFGASLFL-----LSLTVFSI 204
          :|||
Db    170 -----ETLFALPAASEPVIPSSAEKIMDLKEQPGNTVSSGQEDFPSVLFETAASL 219

Qy    205 VSVTAYIALALLSVTISFRI--YKGVIAIQKSDEG-----HPF--R 242
          | : || :||: | : | : || :|||
Db    220 PS-----LSPLS-TVSFKEHGYLGNLSAV-ASTEGTIEETLNEASRELPERATNPFVNR 271

Qy    243 AYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLV-----DDLVDSLKF-----A 290
          | | : : | | : : : :| :|||
Db    272 ESAEFSVLEYSEMGSFNGSPKGESAMLVENTKEEVIVRSKDKEDLVCSAALHNPQESPA 331

Qy    291 VLMWVFTYVGA-----LFNGLTLLILALISLFSIPV-----IYE 324
          | | | :|| : : :| ||
Db    332 TLTKVVKEDGVMSPEKTMDFNEMKMSVVA-----PVREEYADFKPFEQAWEVKDTYE 384

Qy    325 RHQVQIDHYLGLANKSVKDAMAKIQAKI 352
          : : | | | : : :
Db    385 -----GSRDVLAAARANMESKV 400

```

RESULT 14

Q9GM33

ID Q9GM33 PRELIMINARY; PRT; 179 AA.
 AC Q9GM33;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AB049853; BAB16739.1; -.
 DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 179 AA; 19949 MW; 5F8CD4383FEE9E02 CRC64;

Query Match 46.3%; Score 844; DB 6; Length 179;
 Best Local Similarity 95.5%; Pred. No. 8.9e-28;
 Matches 171; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 182 IKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHFP 241
 :|||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYRGVIAIQKSDEGHFP 60
 Qy 242 RAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGA 301
 |||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 RAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLRFVLMWVFTYVGA 120
 Qy 302 LFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
 |||||||||||||||||:||||||| |||||||||||:|||||||||||||||
 Db 121 LFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 179

RESULT 15

Q8K0T0

ID Q8K0T0 PRELIMINARY; PRT; 780 AA.
 AC Q8K0T0;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to reticulon 1 (Hypothetical protein).
 GN RTN1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RA Strausberg R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

Db 317 TVTVSEPEDDSPGSVTPPSSGTEPSAAESQGKGSVSEDELIAAIKEAKGLSYETTESPRP 376

Qy 53 -----ELEVLERKPAAGLSA----- 67
|:|:: | | |

Db 377 VGQVADKPKTKTRSGLPTIPSPLDQEASSAESGDSEIELVSEDPMASEDALPSGYVSFGH 436

Qy 68 -AAVPPAAAAPLLDFS-----SDSVPPAPRGPL 94
: ||: |:| : :| | | | |

Db 437 VSGPPSPASPSIQYSILREEREAEELDSELIIESCDASSASEESPKREQDS-PPMKPGAL 495

Qy 95 PA-----APPAAPERQPSWERSPAAPAP--SLPPAAAVLPSKLPEDDEPPARPPPPPPAG 147
| | | | | : | | | | | | | : | : | |

Db 496 DAIREETGSRATEERAPS-HQGPVEPDPMLSFAPAAALQSRPEPSSGDGASVPEPPRSQQ 554

Qy 148 ASPLAEPAAPPSTPAA---PKRRGSG-----SSVVDLLYWRDIKKTGVVFGASLF 194
| | : :| | | | | : | | | | | | | | | |

Db 555 QKPEEEAVSSSQSPTATEIPGPLGSGLMPPLPFFNKQKAIDLLYWRDIKQTGIVFGSFL 614

Qy 195 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEE 254
| | | | | : | | | | | | | | | | | | | | | | | | | | | | |

Db 615 LLFSLTQFSVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQE 674

Qy 255 LVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALI 314
: | | : : | | | : | | | | | | | | | | | | | | | | | | | | | |

Db 675 QIQKYTDCLQLYVNSTLKELRRLFLVQDLVDLSLKFAVLMWLLTYVGALFNGLTLLIMAVV 734

Qy 315 SLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRRAD 360
| : | : | : | : | | | | | : : | | | | | | | | |

Db 735 SMFTLPVVYVKHQAQVDQYLGLVRTHINTVVAKIQAKIPGAKRHAE 780

Search completed: September 29, 2004, 18:54:39
Job time : 81.3939 secs

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:21:11 ; Search time 14.9792 Seconds
(without alignments)
1251.419 Million cell updates/sec

Title: SEQ2_1-171_975-1163
Perfect score: 1823
Sequence: 1 MEDIDQSSLVSSSTDSPRP.....VKDAMAKIQAKIPGLKRKAD 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	1732.7	95.0	1163	1	RTN4_RAT	Q9jk11 rattus norv
2	1518.4	83.3	1192	1	RTN4_HUMAN	Q9nqc3 homo sapien
3	914	50.1	199	1	RTN4_MOUSE	Q99p72 mus musculu
4	826.3	45.3	776	1	RTN1_HUMAN	Q16799 homo sapien
5	807.4	44.3	777	1	RTN1_RAT	Q64548 rattus norv
6	659.9	36.2	237	1	RTN3_MOUSE	Q9es97 mus musculu
7	654.6	35.9	236	1	RTN3_HUMAN	O95197 homo sapien
8	585.3	32.1	545	1	RTN2_HUMAN	O75298 homo sapien
9	555.3	30.5	471	1	RTN2_MOUSE	O70622 mus musculu
10	272.1	14.9	786	1	PRGR_CHICK	P07812 gallus gall
11	259	14.2	1059	1	CAPU_DROME	Q24120 drosophila
12	257.6	14.1	2715	1	MLL4_HUMAN	Q9umn6 homo sapien
13	257.5	14.1	1206	1	FM14_MOUSE	Q05859 mus musculu
14	256.4	14.1	3530	1	MY15_HUMAN	Q9ukn7 homo sapien
15	254.9	14.0	1468	1	FMN1_MOUSE	Q05860 mus musculu
16	253.9	13.9	1790	1	SEPA_EMENI	P78621 emericella
17	253.7	13.9	1157	1	BBC1_YEAST	P47068 saccharomyc

18	253.2	13.9	1255	1	DIA1_MOUSE	O08808	mus musculu
19	252.5	13.9	1248	1	DIA1_HUMAN	O60610	homo sapien
20	248.6	13.6	555	1	GP1_CHLRE	Q9fpq6	chlamydomon
21	246.4	13.5	1657	1	STNB_CAEEL	P90761	caenorhabdi
22	246	13.5	1522	1	PST1_SCHPO	Q09750	schizosacch
23	245.2	13.5	3511	1	MY15_MOUSE	Q9qzz4	mus musculu
24	244.7	13.4	465	1	FXD1_HUMAN	Q16676	homo sapien
25	242.7	13.3	1533	1	Y586_HUMAN	Q9bvv6	homo sapien
26	242.1	13.3	1375	1	BNR1_YEAST	P40450	saccharomyc
27	241.1	13.2	775	1	ICP0_HSV11	P08393	herpes simp
28	240.9	13.2	980	1	FTKL_DEIRA	Q9rx5	deinococcus
29	240.9	13.2	1427	1	ZFH2_HUMAN	Q9c0a1	homo sapien
30	240.6	13.2	933	1	PRGR_HUMAN	P06401	homo sapien
31	240.5	13.2	1446	1	IE18_PRVKA	P33479	pseudorabie
32	240.2	13.2	1164	1	FHOS_HUMAN	Q9y613	homo sapien
33	238.2	13.1	2517	1	NCR2_HUMAN	Q9y618	h nuclear r
34	238	13.1	3149	1	TEGU_EBV	P03186	epstein-bar
35	237.5	13.0	830	1	JIP2_MOUSE	Q9ere9	mus musculu
36	237.1	13.0	902	1	IF2_BRAJA	Q89wa9	bradyrhizob
37	236.6	13.0	2167	1	SHK1_RAT	Q9wv48	rattus norv
38	236.5	13.0	1395	1	IF4G_HUMAN	Q04637	homo sapien
39	236.3	13.0	1402	1	IF4G_RABIT	P41110	oryctolagus
40	236.3	13.0	2142	1	BAT2_HUMAN	P48634	homo sapien
41	236	12.9	1132	1	BAT3_HUMAN	P46379	homo sapien
42	236	12.9	1332	1	M4K6_HUMAN	Q8n4c8	homo sapien
43	235.5	12.9	925	1	IF2_GLOVI	Q7nh85	gloeobacter
44	235.5	12.9	1461	1	IE18_PRVIF	P11675	pseudorabie
45	234.9	12.9	824	1	JIP2_HUMAN	Q13387	homo sapien

ALIGNMENTS

RESULT 1

RTN4_RAT

ID RTN4_RAT STANDARD; PRT; 1163 AA.
AC Q9JK11; Q9JK10; Q9R0D9; Q9WUE9; Q9WUF0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)
DE (Glut4 vesicle 20 kDa protein).
GN RTN4 OR NOGO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3), AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Adipocyte;
RX MEDLINE=99249816; PubMed=10231557;
RA Morris N.J., Ross S.A., Neveu J.M., Lane W.S., Lienhard G.E.;
RT "Cloning and characterization of a 22 kDa protein from rat adipocytes:
RT a new member of the reticulon family."
RL Biochim. Biophys. Acta 1450:68-76(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).

RX MEDLINE=20129258; PubMed=10667796;
 RA Chen M.S., Huber A.B., Van der Haar M.E., Frank M., Schnell L.,
 RA Spillmann A.A., Christ F., Schwab M.E.;
 RT "Nogo-A is a myelin-associated neurite outgrowth inhibitor and an
 RT antigen for monoclonal antibody IN-1.";
 RL Nature 403:434-439(2000).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
 RC STRAIN=Wistar Kyoto; TISSUE=Vascular smooth muscle;
 RA Ito T., Schwartz S.M.;
 RT "Cloning of a member of the reticulon gene family in rat: one of two
 RT minor splice variants.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP FUNCTION.
 RX MEDLINE=22033691; PubMed=12037567;
 RA GrandPre T., Li S., Strittmatter S.M.;
 RT "Nogo-66 receptor antagonist peptide promotes axonal regeneration.";
 RL Nature 417:547-551(2002).
 CC -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help
 CC block the regeneration of the nervous central system in adults (By
 CC similarity).
 CC -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2 (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Anchored to the
 CC membrane of the endoplasmic reticulum through 2 putative
 CC transmembrane domains (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=1; Synonyms=Nogo-A, NI-220-250;
 CC IsoId=Q9JK11-1; Sequence=Displayed;
 CC Name=2; Synonyms=Nogo-B, Foocen-M1;
 CC IsoId=Q9JK11-2; Sequence=VSP_005658;
 CC Name=3; Synonyms=Nogo-C, VP20;
 CC IsoId=Q9JK11-3; Sequence=VSP_005656, VSP_005657;
 CC Name=4; Synonyms=Foocen-M2;
 CC IsoId=Q9JK11-4; Sequence=VSP_005659;
 CC -!- TISSUE SPECIFICITY: Isoforms 1, 2 and 3 are present in optic
 CC nerve, spinal cord and cerebral cortex. Isoforms 1 and 2 are
 CC present in dorsal root ganglion, sciatic nerve and PC12 cells
 CC after longer exposure. Isoforms 2 and 3 are detected in kidney,
 CC cartilage, skin, lung and spleen. Isoform 3 is expressed at high
 CC level in skeletal muscle. In adult animals isoform 1 is expressed
 CC mainly in the nervous system.
 CC -!- SIMILARITY: Contains 1 reticulon domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF051335; AAF01564.1; -.
 DR EMBL; AJ242961; CAB71027.1; -.
 DR EMBL; AJ242962; CAB71028.1; -.


```

DR      EMBL; AJ242963; CAB71029.1; -.
DR      EMBL; AF132045; AAD31019.1; -.
DR      EMBL; AF132046; AAD31020.1; -.
DR      GO; GO:0030176; C:integral to endoplasmic reticulum membrane; IDA.
DR      GO; GO:0005635; C:nuclear membrane; ISS.
DR      GO; GO:0005515; F:protein binding; ISS.
DR      GO; GO:0019987; P:negative regulation of anti-apoptosis; ISS.
DR      GO; GO:0030517; P:negative regulation of axon extension; ISS.
DR      InterPro; IPR003388; Reticulon.
DR      Pfam; PF02453; Reticulon; 1.
DR      PROSITE; PS50845; RETICULON; 1.
KW      Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT      DOMAIN          1          989          CYTOPLASMIC (Potential).
FT      TRANSMEM        990        1010          POTENTIAL.
FT      DOMAIN          1011        1104          LUMENAL (Potential).
FT      TRANSMEM        1105        1125          POTENTIAL.
FT      DOMAIN          1126        1163          CYTOPLASMIC (Potential).
FT      DOMAIN          976        1163          RETICULON.
FT      DOMAIN          33          46          POLY-GLU.
FT      DOMAIN          73          76          POLY-ALA.
FT      DOMAIN          140        145          POLY-PRO.
FT      VARSPLIC        1          964          Missing (in isoform 3).
FT                                          /FTId=VSP_005656.
FT      VARSPLIC        965        975          AVLSAELSKTS -> MDGQKKHWKDK (in isoform
FT                                          3).
FT                                          /FTId=VSP_005657.
FT      VARSPLIC        173        975          Missing (in isoform 2).
FT                                          /FTId=VSP_005658.
FT      VARSPLIC        192        975          Missing (in isoform 4).
FT                                          /FTId=VSP_005659.
FT      CONFLICT        1130        1131          MISSING (IN REF. 3; AAD31020).
SQ      SEQUENCE        1163 AA; 126386 MW; 8CB894B09E94F0B6 CRC64;

```

Qy	172	-----	171
Db	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQMSVVA	360
Qy	172	-----	171
Db	361	PVREEYADFKPFQAWVEVKDITYEGSRDVLAAARANVESKVDRKCLEDSEGR	420
Qy	172	-----	171
Db	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLEDHTSENKTDEKKIEERKA	480
Qy	172	-----	171
Db	481	QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Qy	172	-----	171
Db	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Qy	172	-----	171
Db	601	LPSAGASVVQPSVSPLEAPPPVSYSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Qy	172	-----	171
Db	661	NAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Qy	172	-----	171
Db	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Qy	172	-----	171
Db	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Qy	172	-----	171
Db	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Qy	172	-----	171
Db	901	DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDTEKED	960
Qy	172	-----SVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	217
Db	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Qy	218	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	277
Db	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Qy	278	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	337
Db	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	1140
Qy	338	NKSVKDAMAKIQAKIPGLKRKAD	360

|||||
Db 1141 NKSVDAMAKIQAKIPGLKRRKAD 1163

RESULT 2

RTN4_HUMAN

ID RTN4_HUMAN STANDARD; PRT; 1192 AA.
AC Q9NQC3; O94962; Q9BXG5; Q9H212; Q9H3I3; Q9UQ42; Q9Y293; Q9Y2Y7;
AC Q9Y5U6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)
DE (Neuroendocrine-specific protein) (NSP) (Neuroendocrine specific
DE protein C homolog) (RTN-x) (Reticulon 5) (My043 protein).
GN RTN4 OR NOGO OR ASY OR KIAA0886.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX MEDLINE=20129242; PubMed=10667780;
RA Prinjha R., Moore S.E., Vinson M., Blake S., Morrow R., Christie G.,
RA Michalovich D., Simmons D.L., Walsh F.S.;
RT "Inhibitor of neurite outgrowth in humans.";
RL Nature 403:383-384(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=21010696; PubMed=11126360;
RA Tagami S., Eguchi Y., Kinoshita M., Takeda M., Tsujimoto Y.;
RT "A novel protein, RTN-XS, interacts with both Bcl-XL and Bcl-2 on
RT endoplasmic reticulum and reduces their anti-apoptotic activity.";
RL Oncogene 19:5736-5746(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX MEDLINE=20237542; PubMed=10773680;
RA Yang J., Yu L., Bi A.D., Zhao S.-Y.;
RT "Assignment of the human reticulon 4 gene (RTN4) to chromosome
RT 2p14-->2p13 by radiation hybrid mapping.";
RL Cytogenet. Cell Genet. 88:101-102(2000).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RA Jin W.-L., Ju G.;
RT "Developmentally-regulated alternative splicing in a novel Nogo-A.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RC TISSUE=Placenta, and Skeletal muscle;
RA Ito T., Schwartz S.M.;
RT "Cloning of a member of the reticulon gene family in human.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Fibroblast;
RA Yutsudo M.;

RT "Isolation of a cell death-inducing gene.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Pituitary;
 RA Song H., Peng Y., Zhou J., Huang Q., Dai M., Mao Y.M., Yu Y., Xu X.,
 RA Luo B., Hu R., Chen J.;
 RT "Human neuroendocrine-specific protein C (NSP) homolog gene.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
 RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
 RA Yu J., Han L.H.;
 RT "Novel human cDNA clone with function of inhibiting cancer cell
 RT growth.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RX MEDLINE=99156230; PubMed=10048485;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 5:355-364(1998).
 RN [10]
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
 RC TISSUE=Brain, Ovary, Pancreas, Placenta, and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [11]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RX MEDLINE=20499367; PubMed=11042152;
 RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
 RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
 RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;

RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells.";
RL Genome Res. 10:1546-1560(2000).
RN [12]
RP SEQUENCE OF 482-1192 FROM N.A. (ISOFORM 1/4).
RC TISSUE=Brain;
RA Mao Y.M., Xie Y., Zheng Z.H.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN [13]
RP SEQUENCE OF 186-1192 FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RA Sha J.H., Zhou Z.M., Li J.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN [14]
RP TOPOLOGY.
RC TISSUE=Brain;
RX MEDLINE=20129259; PubMed=10667797;
RA GrandPre T., Nakamura F., Vartanian T., Strittmatter S.M.;
RT "Identification of the Nogo inhibitor of axon regeneration as a
RT Reticulon protein.";
RL Nature 403:439-444(2000).
RN [15]
RP FUNCTION.
RC TISSUE=Brain;
RX MEDLINE=21069055; PubMed=11201742;
RA Fournier A.E., Grandpre T., Strittmatter S.M.;
RT "Identification of a receptor mediating Nogo-66 inhibition of axonal
RT regeneration.";
RL Nature 409:341-346(2001).
RN [16]
RP REVIEW.
RX MEDLINE=21888956; PubMed=11891768;
RA Ng C.E.L., Tang B.L.;
RT "Nogos and the Nogo-66 receptor: factors inhibiting CNS neuron
RT regeneration.";
RL J. Neurosci. Res. 67:559-565(2002).
CC -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help
CC block the regeneration of the nervous central system in adults.
CC Isoform 2 reduces the anti-apoptotic activity of Bcl-xl and Bcl-2.
CC This is likely consecutive to their change in subcellular
CC location, from the mitochondria to the endoplasmic reticulum,
CC after binding and sequestration.
CC -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum. Anchored to the membrane of the endoplasmic reticulum
CC through 2 putative transmembrane domains.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1; Synonyms=RTN 4A, Nogo-A, RTN-xL;
CC IsoId=Q9NQC3-1; Sequence=Displayed;
CC Name=2; Synonyms=RTN 4B, Nogo-B, RTN-xS, Foocen-M;
CC IsoId=Q9NQC3-2; Sequence=VSP_005655;
CC Name=3; Synonyms=RTN 4C, Nogo-C, Foocen-S;
CC IsoId=Q9NQC3-3; Sequence=VSP_005652, VSP_005653;
CC Name=4;
CC IsoId=Q9NQC3-4; Sequence=VSP_005654;

Db	239	SPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS	298
Qy	171	-----	170
Db	299	EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLVKEDE	358
Qy	171	-----	170
Db	359	VVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLES	418
Qy	171	-----	170
Db	419	KVDKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATN	478
Qy	171	-----	170
Db	479	IFPLLGDPTSENKTDEKKIEEKKQAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTQV	538
Qy	171	-----	170
Db	539	TEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCP	598
Qy	171	-----	170
Db	599	SFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPP	658
Qy	171	-----	170
Db	659	YEEAMSVSLKKVSGIKEEIKEPENINAALQTEAPYISIIACDLIKETKLSAEPAPDFSDY	718
Qy	171	-----	170
Db	719	SEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIQDVPQKQDETVMVLVKESLTETSFES	778
Qy	171	-----	170
Db	779	MIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTA	838
Qy	171	-----	170
Db	839	VYSNDDLFIKSEAQIRETETFSDDSSPIEIIDEFPTLISSKTDSESKLAREYTDLEVSHKS	898
Qy	171	-----	170
Db	899	EIANAPDGAGSLPCTELPHDLNLIQPKVEEKISFSDDFSKNGSATSQVLLLPPDVSA	958
Qy	171	-----SSVVDLLYWRDIKKTG	186
		:	
Db	959	ATQAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTG	1018
Qy	187	VVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPPFRAYLE	246
Db	1019	VVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPPFRAYLE	1078
Qy	247	SEVAISEELVQKYSNSALGHVNSTIKELRRFLVDDLVDLSLKFVLMWVFTYVGALFNGL	306

Db 1079 SEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGFALFNGL 1138

QY 307 TLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
 |||||:||||| |||||:|||||:

Db 1139 TLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192

RESULT 3

RTN4_MOUSE

ID RTN4_MOUSE STANDARD; PRT; 199 AA.

AC Q99P72; Q9CTE3;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein).

GN RTN4 OR NOGO.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=3T3-L1; TISSUE=Adipocyte;

RA Coulson A.C., Craggs P.D., Morris N.J.;

RT "Mouse vp20/RTN4C cDNA.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.

RN [2]

RP SEQUENCE OF 170-199 FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

CC !- FUNCTION: Potent neurite outgrowth inhibitor which may also help

CC block the regeneration of the nervous central system in adults (By

CC similarity).

CC !- SUBUNIT: Binds to RTN4R. Interacts with Bcl-x1 and Bcl-2 (By

CC similarity).

CC !- SUBCELLULAR LOCATION: Integral membrane protein. Anchored to the

CC membrane of the endoplasmic reticulum through 2 putative


```

CC      transmembrane domains (By similarity).
CC      -!- ALTERNATIVE PRODUCTS:
CC          Event=Alternative splicing; Named isoforms=1;
CC          Comment=A number of isoforms may be produced;
CC          Name=1;
CC          IsoId=Q99P72-1; Sequence=Displayed;
CC      -!- SIMILARITY: Contains 1 reticulon domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF326337; AAK08076.1; -.
DR      EMBL; AK003859; -; NOT_ANNOTATED_CDS.
DR      MGD; MGI:1915835; Rtn4.
DR      GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR      GO; GO:0030176; C:integral to endoplasmic reticulum membrane; ISS.
DR      GO; GO:0005635; C:nuclear membrane; ISS.
DR      GO; GO:0005515; F:protein binding; ISS.
DR      GO; GO:0019987; P:negative regulation of anti-apoptosis; ISS.
DR      GO; GO:0030517; P:negative regulation of axon extension; ISS.
DR      GO; GO:0007399; P:neurogenesis; IDA.
DR      InterPro; IPR003388; Reticulon.
DR      Pfam; PF02453; Reticulon; 1.
DR      PROSITE; PS50845; RETICULON; 1.
KW      Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT      DOMAIN      1      25      CYTOPLASMIC (Potential).
FT      TRANSMEM      26      50      POTENTIAL.
FT      DOMAIN      51      137      LUMENAL (Potential).
FT      TRANSMEM      138      162      POTENTIAL.
FT      DOMAIN      163      199      CYTOPLASMIC (Potential).
FT      DOMAIN      12      199      RETICULON.
SQ      SEQUENCE      199 AA; 22466 MW; 07BE5D580059ED9C CRC64;

Query Match      50.1%; Score 914; DB 1; Length 199;
Best Local Similarity 83.3%; Pred. No. 1.7e-24;
Matches 190; Conservative 3; Mismatches 5; Indels 30; Gaps 2;

Qy      133 DDEPPARPPPPPPAGASPLAEPAPSTPAAPKRRGSGSSVVDLLYWRDIKKTGVVFGAS 192
      ||:                                     |:|  |||||
Db      2 DDQ-----KRWK-DKVVDLLYWRDIKKTGVVFGAS 31

Qy      193 LFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAIS 252
      |||||
Db      32 LFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAIS 91

Qy      253 EELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGALFNGLTLLILA 312
      |||||
Db      92 EELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGALFNGLTLLILA 151

Qy      313 LISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRRKAD 360
      |||||
Db      152 LISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRRKAE 199

```

RESULT 4

RTN1_HUMAN

ID RTN1_HUMAN STANDARD; PRT; 776 AA.
AC Q16799; Q16800; Q16801;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Reticulon 1 (Neuroendocrine-specific protein).
GN RTN1 OR NSP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS RTN1-A; RTN1-B AND RTN1-C).
RC TISSUE=Lung carcinoma;
RX MEDLINE=93293865; PubMed=7685762;
RA Roebroek A.J.M., Van de Velde H.J.K., Van Bokhoven A., Broers J.L.V.,
RA Ramaekers F.C.S., Van de Ven W.J.M.;
RT "Cloning and expression of alternative transcripts of a novel
RT neuroendocrine-specific gene and identification of its 135-kDa
RT translational product.";
RL J. Biol. Chem. 268:13439-13447(1993).
RN [2]
RP ALTERNATIVE SPLICING.
RX MEDLINE=96429995; PubMed=8833145;
RA Roebroek A.J.M., Ayoubi T.A.Y., Van de Velde H.J.K.,
RA Schoenmakers E.F.P.M., Pauli I.G.L., Van de Ven W.J.M.;
RT "Genomic organization of the human NSP gene, prototype of a novel gene
RT family encoding reticulons.";
RL Genomics 32:191-199(1996).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=98228245; PubMed=9560466;
RA Hens J., Nuydens R., Geerts H., Senden N.H., Van de Ven W.J.M.,
RA Roebroek A.J., van de Velde H.J.K., Ramaekers F.C., Broers J.L.;
RT "Neuronal differentiation is accompanied by NSP-C expression.";
RL Cell Tissue Res. 292:229-237(1998).
CC -!- FUNCTION: May be involved in neuroendocrine secretion or in
CC membrane trafficking in neuroendocrine cells.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=RTN1-A; Synonyms=NSP-A;
CC IsoId=Q16799-1; Sequence=Displayed;
CC Name=RTN1-B; Synonyms=NSP-B;
CC IsoId=Q16799-2; Sequence=VSP_005644;
CC Name=RTN1-C; Synonyms=NSP-C;
CC IsoId=Q16799-3; Sequence=VSP_005645, VSP_005646;
CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURAL AND NEUROENDOCRINE TISSUES
CC AND CELL CULTURES DERIVED THEREFROM. EXPRESSION OF ISOFORM RTN1-C
CC IS STRONGLY CORRELATED WITH NEURONAL DIFFERENTIATION.
CC -!- PTM: Isoforms RTN1-A and RTN1-B are phosphorylated.
CC -!- SIMILARITY: Contains 1 reticulon domain.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L10333; AAA59950.1; -.
 DR EMBL; L10334; AAA59951.1; -.
 DR EMBL; L10335; AAA59952.1; -.
 DR PIR; A46583; A46583.
 DR PIR; I60904; I60904.
 DR Genew; HGNC:10467; RTN1.
 DR MIM; 600865; -.
 DR GO; GO:0030176; C:integral to endoplasmic reticulum membrane; TAS.
 DR GO; GO:0004871; F:signal transducer activity; NAS.
 DR GO; GO:0030182; P:neuron differentiation; TAS.
 DR GO; GO:0007165; P:signal transduction; NAS.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 KW Endoplasmic reticulum; Alternative splicing; Transmembrane;
 KW Phosphorylation.
 FT TRANSMEM 603 623 POTENTIAL.
 FT TRANSMEM 726 746 POTENTIAL.
 FT DOMAIN 589 776 RETICULON.
 FT DOMAIN 609 612 POLY-LEU.
 FT VARSPLIC 1 420 Missing (in isoform RTN1-B).
 FT /FTid=VSP_005644.
 FT VARSPLIC 1 568 Missing (in isoform RTN1-C).
 FT /FTid=VSP_005645.
 FT VARSPLIC 569 588 GPGPLGPGAPPPLLEFLNKQK -> MQATADSTKMDCVWSNW
 FT KSQ (in isoform RTN1-C).
 FT /FTid=VSP_005646.
 SQ SEQUENCE 776 AA; 83617 MW; CA5B6232353096FE CRC64;

Query Match 45.3%; Score 826.3; DB 1; Length 776;
 Best Local Similarity 27.7%; Pred. No. 4.2e-20;
 Matches 195; Conservative 52; Mismatches 91; Indels 367; Gaps 17;

Qy 10 VSSSTDS-----PPR----- 19
 |||: | ||:
 Db 85 VSSAMDHTFSTTSKDGEGSCYTSLISDICYPPQEDSTYFTGILQKENGHVITISESPEELG 144
 Qy 20 -PPP-----AFKYQFVTEPED 34
 | | |::| |::
 Db 145 TPGPSLPDVPGIESRGLFSSDSGIEMTPAESTEVNKILADPLDQMKAAYKYIDITRPEE 204
 Qy 35 EEDEEEEEDEEED-----DED----- 50
 :::| || |:
 Db 205 VKHQEQHHPELEDKDLDFKNKDTDISIKPEGVREPDKPAPVEGKI IKDHLLLEESTFAPYI 264
 Qy 51 -----LEELE----- 55
 | |:
 Db 265 DDLSEEQRRAPQITTPVKITLTEIEPSVETTTQEKTPKQDICKPSPDTVPTVTVSEPE 324

QY 56 -----VLERKPAAGLS----- 66
: | | | |
Db 325 DDSPGSITPPSSGTEPSAAESQGKGSISEDELITAIKEAKGLSYETAENPRPVGQLADRP 384
QY 67 ---AAAVPPAAAAPLLDFSSDSV-----PPA 89
| : | | : | | | : | | :
Db 385 EVKARSGPPTIPSP-LDHEASSAESGDSEIELVSEDPMMAEDALPSGYVSFGHVGGPPPS 443
QY 90 PRGPL-----PAAPPA----- 100
| | | | |
Db 444 PASPSIQYSILREEREAELDSELIIESCDASSASEESPKREQDSPMKPSALDAIREETG 503
QY 101 --APERQPS-----WERSPAAPAPSLPPA-----AAVLPSKLPEDD-----E 135
| | | | : : | | | | : | | | : | :
Db 504 VRAERAPSRRLAEPGSFLDYPSTEPQGPPELPPGDGALEPETPMLPRK-PEEDSSSNQ 562
QY 136 PPARPPPPPPAGASPLAEPAPPSTPAAPKRRGSGSSVVDLLYWRDIKKTGVVFGASLFL 195
| | | | | | | | : : : | | | | : | : | : | |
Db 563 SPAATKGPGLG-----PGAPPPLLFLNKQK-----AIDLLYWRDIKQTGIVFGSFLLL 611
QY 196 LLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEEL 255
| | | | | : | | | | | | | | | : | : | : | | | : : | : |
Db 612 LFSLTQFSVSVVAYIALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQ 671
QY 256 VQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALIS 315
: | : : : : : | : | : | : | : | : | : | : | : | : : | : |
Db 672 IQKYTDCLQFYVNSTLKELRRLFLVQDLVDLSLKFAVLMWLLTYVGALFNGLTLLLMVAVS 731
QY 316 LFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 360
: | : : | : | | | | | : : | : | : | | : | :
Db 732 MFTLPVVYVKHQAQIDQYLGLVLRTHINAVVAKIQAKIPGAKRHAE 776

RESULT 5

RTN1_RAT

ID RTN1_RAT STANDARD; PRT; 777 AA.
AC Q64548; Q64547;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Reticulon 1 (Neuroendocrine-specific protein) (S-rex).
GN RTN1 OR NSP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS RTN1-B AND RTN1-S).
RC STRAIN=Wistar; TISSUE=Brain cortex;
RX MEDLINE=96386034; PubMed=8793864;
RA Baka I.D., Ninkina N.N., Pinon L.G.P., Adu J., Davies A.M.,
RA Georgiev G.P., Buchman V.L.;
RT "Intracellular compartmentalization of two differentially spliced s-
RT rex/NSP mRNAs in neurons."
RL Mol. Cell. Neurosci. 7:289-303(1996).
CC -!- FUNCTION: May be involved in neuroendocrine secretion or in
CC membrane trafficking in neuroendocrine cells.

```

CC  -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane (By
CC      similarity).
CC  -!- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
CC      Name=RTN1-B; Synonyms=S-RexB;
CC      IsoId=Q64548-1; Sequence=Displayed;
CC      Name=RTN1-S; Synonyms=S-RexS;
CC      IsoId=Q64548-2; Sequence=VSP_005647, VSP_005648;
CC  -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CENTRAL AND
CC      PERIPHERAL NERVOUS SYSTEM OF NEWBORN AND ADULT RATS. LOW LEVELS
CC      HAVE BEEN ALSO DETECTED IN HEART, ADRENAL GLAND AND SPLEEN.
CC      EXPRESSION OF ISOFORM RTN1-B IS RESTRICTED TO PARTICULAR NEURONAL
CC      TYPES.
CC  -!- DEVELOPMENTAL STAGE: DETECTED ON EMBRYONIC DAY E10 IN THE
CC      HINDBRAIN AND IN E11 IN THE FOREBRAIN. DURING THE NEXT 3 EMBRYONIC
CC      DAYS THE LEVELS OF S-REXS INCREASES AND REMAINS STABLE AT E13 IN
CC      THE HINDBRAIN AND AT E14 IN THE FOREBRAIN. THE LEVELS OF S-REXB
CC      DOES NOT CHANGE AS SIGNIFICANTLY DURING DEVELOPMENT OF THE
CC      HINDBRAIN.
CC  -!- SIMILARITY: Contains 1 reticulon domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U17604; AAC53046.1; -.
DR  EMBL; U17603; AAC53045.1; -.
DR  InterPro; IPR003388; Reticulon.
DR  Pfam; PF02453; Reticulon; 1.
DR  PROSITE; PS50845; RETICULON; 1.
KW  Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT  TRANSMEM      604      624      POTENTIAL.
FT  TRANSMEM      727      747      POTENTIAL.
FT  DOMAIN        590      777      RETICULON.
FT  DOMAIN        610      613      POLY-LEU.
FT  VARSPLIC       1      569      Missing (in isoform RTN1-S).
FT                                     /FTId=VSP_005647.
FT  VARSPLIC      570      589      IPGPLGSDLVPPLPFFNKQK -> MQATADSTKMDCVWSNW
FT                                     KSQ (in isoform RTN1-S).
FT                                     /FTId=VSP_005648.
SQ  SEQUENCE      777 AA;  83001 MW;  AF7479C50F28D0AC CRC64;

```

```

Query Match          44.3%;  Score 807.4;  DB 1;  Length 777;
Best Local Similarity 27.9%;  Pred. No. 1.8e-19;
Matches 189;  Conservative 61;  Mismatches 101;  Indels 326;  Gaps 15;

```

```

Qy      7  SSLVS-----SSTDSPPR---PPP----- 22
      :||:|                ::::| |
Db      104 TSLISDICYPPREDSAYFTGILQKENGHITTSESPEELGTPGPSLPEVPGTEPHGLLSSD 163

Qy      23  -----AFKYQFVTEPEDEEDEEEEE---EDEEED--DE 49
      | || :| |:: : :||:  ||:: | |:
Db      164 SGIEMTPAESTEVNKKILADPLDQMKAEACKYIDITRPQEAQGQEEQSPGLEDKDLDFKDK 223

```

Qy 50 DLE----- 52
 | |
 Db 224 DSEVSTKPEGVHAPNQSPVEGKLIKDNLFEESTFAPYIDELSDEQHRMSLV TAPVKITL 283
 Qy 53 ----- 52
 Db 284 TEIGPPVMTATHETIPEKQDLCLKPSPDTVPTVTVSEPEDDSPGSVTPPSSGTEPSAAES 343
 Qy 53 ----- 52
 Db 344 QGKGSVSEDELIAAIKEAKGLSYETTESPRPVGQAADRPVKARSGLPTIPSSLDQEASS 403
 Qy 53 -----ELEVLERKPAAGLSA-----AAVPPAAAAPLLDfs----- 82
 |:|:: | | | : ||: |:| : :|
 Db 404 AESGDSEIELVSEDPMASEDALPSGYVSFGHVSGPPSPASPSIQYSILREEREAEELDSE 463
 Qy 83 -----SDSVPPAPRGPLPA-----APPAAPERQPSWERSPAAPAPS 118
 | | | | | : | | | | : | | |
 Db 464 LIIESCDASSASEESPKREQDS-PPMKPGVLD AIREETSSRATEERAPS-HQGPVEPDPI 521
 Qy 119 LPPAAAVLPSKLPEDDEPPARPPPPPPAG--ASPLAEPAAPPSTPAAPKRRGS-GSSVV- 174
 | | : | | | | | : : | | : | | : |
 Db 522 LSFTPVTLQSR-PEPSSGDGAPVPEPPKSQQQKPEEEAVSSSQSPAATEIPGPLGSDLVP 580
 Qy 175 -----DLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFR 223
 | | | | | : | | | : | | | | : | | | | | | |
 Db 581 PLPFFNKQKAIDLlyWRDIKQTGIVFGSFLLLLSLTQFSVSVVAYLALAALSATISFR 640
 Qy 224 IYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDL 283
 | | | : | | : | | | | : | : | : | | : | | | | | | | | | |
 Db 641 IYKSVLQAVQKTDEGHPFKAYLELEITLSQEIQKYTDCLQLYVNSTLKE LRRLFLVQDL 700
 Qy 284 VDSLKFAVLMWVFTYVGALENGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKD 343
 | | | | | : | | | | | : | : | : | : | : | | | | :
 Db 701 VDSLKFAVLMWLLTYVGALENGLTLLLMAVSMFTLPVVYVKHQAQVDQYLGLVRTHINT 760
 Qy 344 AMAKIQAKIPGLKRKAD 360
 : | | | | | | | | | | :
 Db 761 VVAKIQAKIPGAKRHAE 777

RESULT 6

RTN3_MOUSE

ID RTN3_MOUSE STANDARD; PRT; 237 AA.

AC Q9ES97;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Reticulon protein 3.

GN RTN3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Huang X., Zhou Y., Qiang H., Yuan J., Qiang B.;
 RT "Cloning and expression profile of a novel mouse cDNA encoding a human
 RT RTN3 homolog.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 CC reticulum (Potential).
 CC -!- SIMILARITY: Contains 1 reticulon domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF195940; AAG31360.1; -.
 DR EMBL; BC014697; AAH14697.1; -.
 DR MGD; MGI:1339970; Rtn3.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 KW Transmembrane; Endoplasmic reticulum.
 FT TRANSMEM 69 89 POTENTIAL.
 FT TRANSMEM 167 187 POTENTIAL.
 FT DOMAIN 49 237 RETICULON.
 SQ SEQUENCE 237 AA; 25428 MW; EB60A0A7AC45F0DE CRC64;

Query Match 36.2%; Score 659.9; DB 1; Length 237;
 Best Local Similarity 49.8%; Pred. No. 9.6e-16;
 Matches 129; Conservative 47; Mismatches 52; Indels 31; Gaps 4;

Qy 109 ERSPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLA-----EPAAPPSTPA 162

```

      | | | :|: ::                :| | | | | | : |
Db      3 ESSAATQSPSVSSSS-----SGAEPSALGGGGGSPGACPALGA 40
Qy      163 APKRRGSGSSVVDLLYWRDIKKTGVVFGASLFLLSLTVFSIVSVTAYIALALLSVTISF 222
      | | | :| | |:| |:| | | | | | | | | | | | | | | |
Db      41 --KSCGSSCAVHDLIFWRDVKKTGFVFGTTLIMLLSLAAFSVISVSYLILALLSVTISF 98
Qy      223 RIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDD 282
      |:| | |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db      99 RVYKSVIQAVQKSEEGHPFKAYLDVDITLSSEAFHNYMNAAMVHVNKALKLIIRLFLVED 158
Qy      283 LVDSLKFVILMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVK 342
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      159 LVDSLKLAVFMWLMTYVGAVFNGITLLILAEILVFSVPPIVYEKYKTQIDHYVGIARDQTK 218
Qy      343 DAMAKIQAKIPGL-KRKAD 360
      : | | | | |:| |:| |:| |:|
Db      219 SIVEKIQAKLPGIAKKKAE 237

```

RESULT 7

RTN3_HUMAN

```

ID   RTN3_HUMAN          STANDARD;          PRT;   236 AA.
AC   O95197;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Reticulon protein 3 (Neuroendocrine-specific protein-like 2) (NSP-like
DE   protein II) (NSPLII).
GN   RTN3 OR NSPL2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Retina;
RX   MEDLINE=99265974; PubMed=10331947;
RA   Moreira E.F., Jaworski C.J., Rodriguez I.R.;
RT   "Cloning of a novel member of the reticulon gene family (RTN3): gene
RT   structure and chromosomal localization to 11q13.";
RL   Genomics 58:73-81(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Huang X., Zhou Y., Du G., Yuan J., Qiang B.;
RT   "Cloning and expression analysis of a cDNA encoding a novel
RT   neuroendocrine-specific protein-like protein 1: NSPL1.";
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, Eye, and Lymph;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```


RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 CC reticulum (Potential).

CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSION IN
 CC BRAIN. THREE TIMES MORE ABUNDANT IN MACULA THAN IN PERIPHERAL
 CC RETINA.

CC -!- SIMILARITY: Contains 1 reticulon domain.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF059524; AAC99319.1; -.

DR EMBL; AF059529; AAD20951.1; -.

DR EMBL; AF059525; AAD20951.1; JOINED.

DR EMBL; AF059526; AAD20951.1; JOINED.

DR EMBL; AF059527; AAD20951.1; JOINED.

DR EMBL; AF059528; AAD20951.1; JOINED.

DR EMBL; AF119297; AAD26810.1; -.

DR EMBL; BC000634; AAH00634.1; -.

DR EMBL; BC010556; AAH10556.1; -.

DR EMBL; BC011394; AAH11394.1; -.

DR EMBL; BC022993; AAH22993.1; -.

DR Genew; HGNC:10469; RTN3.

DR MIM; 604249; -.

DR GO; GO:0005615; C:extracellular space; TAS.

DR InterPro; IPR003388; Reticulon.

DR Pfam; PF02453; Reticulon; 1.

DR PROSITE; PS50845; RETICULON; 1.

KW Transmembrane; Endoplasmic reticulum.

FT TRANSMEM 68 88 POTENTIAL.

FT TRANSMEM 177 197 POTENTIAL.

FT DOMAIN 48 236 RETICULON.

SQ SEQUENCE 236 AA; 25609 MW; DDC6A4544ABCD7B7 CRC64;

Query Match 35.9%; Score 654.6; DB 1; Length 236;

Best Local Similarity 47.8%; Pred. No. 1.4e-15;

Matches 129; Conservative 44; Mismatches 57; Indels 40; Gaps 4;

Qy	96	AAPPAAPE-----RQPSWERSPAAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPL	151
Db	2	AEPSAATQSHSISSSSFGAEPSAPG-----GGGSPG	32
Qy	152	AEPAAPPSTPAAPKRRGSGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYI	211
Db	33	ACPA-----LGTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLIMLLSLAAFSVISVVSYL	86
Qy	212	ALALLSVTISFRIYKGVIAIQKSDGHPFRAYLESEVAISEELVQKYSNSALGHVNSTI	271
Db	87	ILALLSVTISFRIYKSVIQAVQKSEEGHPFKAYLDVDITLSSEAFHNYMNAAMVHINRAL	146
Qy	272	KELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQID	331
Db	147	KLIIRLFLVEDLVDLSLKLAVFMWLMTYVGAVFNGITLLILAELLIFSVPIVYEKYKTQID	206
Qy	332	HYLGLANKSVKDAMAKIQAKIPGL-KRKAD	360
Db	207	HYVGIARDOTKSIVEKIQAKLPGIAKKKAE	236

RTN2 HUMAN

```

AC      075298; 060509;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Reticulon protein 2 (Neuroendocrine-specific protein-like 1) (NSP-like
DE      protein 1) (NSPLI).
GN      RTN2 OR NSPL1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC      TISSUE=Lung carcinoma;
RX      MEDLINE=98360096; PubMed=9693037;
RA      Roebroek A.J.M., Contreras B., Pauli I.G.L., Van de Ven W.J.M.;
RT      "cDNA cloning, genomic organization, and expression of the human RTN2
RT      gene, a member of a gene family encoding reticulons.";
RL      Genomics 51:98-106(1998).
RN      [2]
RP      SEQUENCE OF 108-545 FROM N.A. (ISOFORM RTN2-B).
RC      TISSUE=Brain;
RX      MEDLINE=98191726; PubMed=9530622;
RA      Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;
RT      "Molecular cloning of a novel mouse gene with predominant muscle and
RT      neural expression.";
RL      Mamm. Genome 9:274-282(1998).
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC      reticulum (Potential).
CC      -!- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
CC      Name=RTN2-A;
CC      IsoId=075298-1; Sequence=Displayed;

```

CC Note=Isoform RTN2-C is produced by alternative initiation at
 CC Met-341 of isoform RTN2-A;
 CC Name=RTN2-B;
 CC IsoId=O75298-2; Sequence=VSP_005649;
 CC Event=Alternative initiation;
 CC Comment=2 isoforms, RTN2-A (shown here) and RTN2-C, are produced
 CC by alternative initiation at Met-1 and Met-341;
 CC -!- TISSUE SPECIFICITY: ISOFORM RTN2-C IS HIGHLY EXPRESSED IN SKELETAL
 CC MUSCLE.
 CC -!- SIMILARITY: Contains 1 reticulon domain.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF004222; AAC32542.1; -.
 DR EMBL; AF004223; AAC32543.1; -.
 DR EMBL; AF004224; AAC32544.1; -.
 DR EMBL; AF038540; AAC14910.1; -.
 DR Genew; HGNC:10468; RTN2.
 DR MIM; 603183; -.
 DR GO; GO:0030176; C:integral to endoplasmic reticulum membrane; NAS.
 DR GO; GO:0004871; F:signal transducer activity; NAS.
 DR GO; GO:0007165; P:signal transduction; NAS.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 KW Endoplasmic reticulum; Alternative splicing; Transmembrane;
 KW Alternative initiation.
 FT CHAIN 1 545 RETICULON PROTEIN 2, ISOFORM RTN2-A.
 FT CHAIN 341 545 RETICULON PROTEIN 2, ISOFORM RTN2-C.
 FT INIT_MET 341 341 FOR ISOFORM RTN2-C.
 FT TRANSMEM 368 388 POTENTIAL.
 FT TRANSMEM 463 483 POTENTIAL.
 FT DOMAIN 345 545 RETICULON.
 FT VARSPLIC 272 344 Missing (in isoform RTN2-B).
 FT /FTId=VSP_005649.
 SQ SEQUENCE 545 AA; 59263 MW; 971FD2F909E1E9E6 CRC64;

Query Match 32.1%; Score 585.3; DB 1; Length 545;
 Best Local Similarity 27.5%; Pred. No. 2.3e-12;
 Matches 154; Conservative 52; Mismatches 117; Indels 237; Gaps 17;

Qy 8 SLVSSSTDSPRRPPPAFKYQFVTE-PEDEED-----EEEEEEDEED----- 47
 | ||: || || | : | | |||||
 Db 16 STASSTPDS-----TEGGNDDSDFRELHTAREFSEEEDEETTSQDWGTPREL 62
 Qy 48 -----DEDLEELEVLERKPAAG 64
 : || : | : | |
 Db 63 TFSYIAFDGVVGSGGRRDSTARRPRPQGRSVSEPRDQHPQPSLGSLESIPSLQSPEPG 122
 Qy 65 LSAAAVPPAAAAAPLLDFSSDSVPPAPR-----GPL 94
 | : ||: | ||

```

Db      123 RRG-----DPDTAPPSERPLEDLRLRLDHLGWVARGTSGSGEDSSTSSSTPL 168
QY      95 PAAPPAAPER-----QPSWERSPAAPAPSL-----PPAAVLPSK-- 129
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      169 EDEEPQEPNRLETGEAGEELDLRLRLAQPS---SPEVLTPQLSPGSGTPQAGTPSPSRSR 225
QY      130 -----LPEDDE----PPARPP-----PPPPAGAS----- 149
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      226 DSNSGPEEPLEEELKQWGPLEREPVRGQCLDSTDQLEFTVEPRLLTAMEWLKTSLLLA 285
QY      150 -----PLAEPAP-----PSTPA-----APKRRG-----SGSSV 173
      | : | : | | | | | | | | | | | | | | | | | | | | | | | |
Db      286 VYKTVPILELSPPLWTAIGWVQRGPTPPTPVLRVLLKWAKSPRSSGVPSLSLGADMGSKV 345
QY      174 VDLLYWDRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQ 233
      | | | | | : : : | | | | | | | | | | | | | | | | | | | | |
Db      346 ADLLYWKDTRTSGVVFTGLMVSLLCLLHFSIVSVAHLALLLLCGTISLRVYRKVLQAVH 405
QY      234 KSDEGHPPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLM 293
      : | : | | | | : : : | : : | : | : | | : | | | | | | | | |
Db      406 RGDGANPFQAYLDVDLTLTREQTERLSHQITSRVVSAATQLRHFFLVEDLVDLSLKALLF 465
QY      294 WVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIP 353
      : : | : | | | | | | | | | | | | | | | | | | | | | | | |
Db      466 YILTFVGAIFNGLTLLILGVIGLFTIPLLYRQHQQAQIDQYVGLVTNQLSHIKAKIRAKIP 525
QY      354 GL-----KRKAD 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      526 GTGALASAAAASVSGSKAKAE 545

```

RESULT 9

RTN2_MOUSE

```

ID   RTN2_MOUSE      STANDARD;          PRT;   471 AA.
AC   O70622; O70620;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Reticulon protein 2 (Neuroendocrine-specific protein-like 1) (NSP-like
DE   protein 1) (NSPL1).
GN   RTN2 OR NSPL1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND TISSUE SPECIFICITY.
RC   STRAIN=FVB/N, and 129/Sv; TISSUE=Cerebellum, and Skeletal muscle;
RX   MEDLINE=98191726; PubMed=9530622;
RA   Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;
RT   "Molecular cloning of a novel mouse gene with predominant muscle and
RT   neural expression.";
RL   Mamm. Genome 9:274-282(1998).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Retina;
RX   MEDLINE=22388257; PubMed=12477932;

```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
 CC (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=Brain;
 CC IsoId=O70622-1; Sequence=Displayed;
 CC Name=2; Synonyms=Muscle;
 CC IsoId=O70622-2; Sequence=VSP_005650, VSP_005651;
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in neural and muscular
 CC tissues.
 CC -!- SIMILARITY: Contains 1 reticulon domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF038537; AAC14906.1; -.
 DR EMBL; AF038537; AAC14907.1; -.
 DR EMBL; AF038538; AAC14908.1; -.
 DR EMBL; AF038539; AAC14909.1; -.
 DR EMBL; AF093624; AAD13195.1; -.
 DR EMBL; BC031370; AAH31370.1; -.
 DR MGD; MGI:107612; Rtn2.
 DR InterPro; IPR003388; Reticulon.
 DR Pfam; PF02453; Reticulon; 1.
 DR PROSITE; PS50845; RETICULON; 1.
 KW Endoplasmic reticulum; Alternative splicing; Transmembrane.
 FT TRANSMEM 295 315 POTENTIAL.
 FT DOMAIN 272 471 RETICULON.
 FT VARSPLIC 1 267 Missing (in isoform 2).
 FT /FTId=VSP_005650.
 FT VARSPLIC 268 271 PLLL -> MGSK (in isoform 2).
 FT /FTId=VSP_005651.

SQ SEQUENCE 471 AA; 51346 MW; 9BBD8F372CF63AD3 CRC64;

Query Match 30.5%; Score 555.3; DB 1; Length 471;
Best Local Similarity 27.0%; Pred. No. 1.7e-11;
Matches 140; Conservative 51; Mismatches 100; Indels 227; Gaps 13;

```
Qy      8 SLVSSSTDSPPRPPPAFKYQFVTEP-----EDEEEEEEEEEDEEDEDLEE 53
      | ||: ||                ||                | ||||| :| |
Db     16 STASSTPDS-----TEGGNDDSDFRELHTAREFSEDEEEETTSQDWGTPRE- 61

Qy     54 LEVLERKPAAGLSAAAVPPAAAAPLLDFS-----SDSV--PPAPRGPLPAAP 98
      | ||                |||| | |:| :|
Db     62 -----LTFSYIAFDGVVSGGRRDSVVRPRPQGRSVSEP 96

Qy     99 PAAPERQ-----PSWERSPAAPAPSLPPAAAVLPSKLPEDDEPPAR--PPPPPPAG 147
      |::                || :||                || | | | |||
Db     97 RDPPQQSGLGDSLESIPSLSQSP-----EPGRRGDPDPVPPA- 133

Qy    148 ASPLAE----- 153
      || |
Db    134 ERPLEELRLRLDQLGWVRSAGSGEDSATSSSTPLENEEPDGLLEASEAGEETNLELRLAQ 193

Qy    154 -----PAAPPS-----TPAAPKRRGSGSS----- 172
      | ||                ||: : : | |
Db    194 SLHLQLEVLTPQLSPSSGTPQAHTPSPQRSQDSNSGPDDEPLLNVEEHWRLLQEPEPITA 253

Qy    173 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALA 214
      | ||||:| : :| || : || | ||||| |:| |
Db    254 QCLDSTDQSEFMLEPLLLVADLLYWKDTRTSGAVFTGLMASLLCLLHFSIVSVA AHLALL 313

Qy    215 LLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKEL 274
      | ||| |:| : |:| : :| :||:| : : :| : : | || || :|
Db    314 GLCATISLRVYRKVLQAVHRGDGTNPFQAYLDMDLTLTREQTERLSQQIASHVVSTATQL 373

Qy    275 RRLFLVDDLVDLSLKFAVLMWVFTYVGFALFNGLTLLILALISLFSIPVIYERHQVQIDHYL 334
      | |||:||||| |:| : :|:|:|:|:|:| : :|:|:|:| :| ||| |:|
Db    374 RHFFLVEDLVDSLKLALLFYILTFTVGAIFNGLTLVILGVVALFTVPLLYRQHQQAQIDQYV 433

Qy    335 GLANKSVKDAMAKIQAKIPGL-----KRKAD 360
      || : |||:||||| | ||:
Db    434 GLVTNQLSHIKAKIRAKIPGTGTLAPTASVSGSKAKAE 471
```

RESULT 10

PRGR_CHICK

ID PRGR_CHICK STANDARD; PRT; 786 AA.
AC P07812; Q90946;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progesterone receptor (PR).
GN PGR OR NR3C3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88166640; PubMed=3443098;
 RA Gronemeyer H., Turcotte B., Quirin-Stricker C., Bocquel M.T.,
 RA Meyer M.E., Krozowski Z., Jeltsch J.M., Lerouge T., Garnier J.M.,
 RA Chambon P.;
 RT "The chicken progesterone receptor: sequence, expression and
 RT functional analysis.";
 RL EMBO J. 6:3985-3994(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91042592; PubMed=3153474;
 RA Conneely O.M., Dobson A.D.W., Tsai M.-J., Beattie W.G., Toft D.O.,
 RA Huckaby C.S., Zarucki T., Schrader W.T., O'Malley B.W.;
 RT "Sequence and expression of a functional chicken progesterone
 RT receptor.";
 RL Mol. Endocrinol. 1:517-525(1987).
 RN [3]
 RP SEQUENCE OF 128-164 FROM N.A.
 RX MEDLINE=86289413; PubMed=2426779;
 RA Conneely O.M., Sullivan W.P., Toft D.O., Birnbaumer M., Cook R.G.,
 RA Maxwell B.L., Zarucki-Schulz T., Greene G.L., Schrader W.T.,
 RA O'Malley B.W.;
 RT "Molecular cloning of the chicken progesterone receptor.";
 RL Science 233:767-770(1986).
 RN [4]
 RP SEQUENCE OF 417-490 FROM N.A.
 RX MEDLINE=86287271; PubMed=2426697;
 RA Jeltsch J.M., Krozowski Z., Quirin-Stricker C., Gronemeyer H.,
 RA Simpson R.J., Garnier J.M., Krust A., Jacob F., Chambon P.;
 RT "Cloning of the chicken progesterone receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5424-5428(1986).
 RN [5]
 RP DIFFERENCE BETWEEN FORM 1 AND FORM 2.
 RX MEDLINE=89340509; PubMed=2760059;
 RA Conneely O.M., Kettelberger D.M., Tsai M.-J., Schrader W.T.,
 RA O'Malley B.W.;
 RT "The chicken progesterone receptor A and B isoforms are products of
 RT an alternate translation initiation event.";
 RL J. Biol. Chem. 264:14062-14064(1989).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORMS A; A'; B AND B').
 RX MEDLINE=90154085; PubMed=2303488;
 RA Jeltsch J.-M., Turcotte B., Garnier J.-M., Lerouge T., Krozowski Z.,
 RA Gronemeyer H., Chambon P.;
 RT "Characterization of multiple mRNAs originating from the chicken
 RT progesterone receptor gene. Evidence for a specific transcript
 RT encoding form A.";
 RL J. Biol. Chem. 265:3967-3974(1990).
 CC -!- FUNCTION: The steroid hormones and their receptors are involved in
 CC the regulation of eukaryotic gene expression and affect cellular
 CC proliferation and differentiation in target tissues.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=A;
 CC IsoId=P07812-1; Sequence=Displayed;

```

CC      Name=A';
CC      IsoId=P07812-2; Sequence=VSP_003708, VSP_003709;
CC      Name=B;
CC      IsoId=P07812-3; Sequence=VSP_003707;
CC      Name=B';
CC      IsoId=P07812-4; Sequence=VSP_003707, VSP_003708, VSP_003709;
CC      -!- TISSUE SPECIFICITY: Oviduct.
CC      -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC              a DNA-binding domain and a C-terminal steroid-binding domain.
CC      -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC              subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Y00092; CAA68282.1; -.
DR      EMBL; M13972; AAA49034.1; -.
DR      EMBL; M37518; AAA49013.1; -.
DR      EMBL; M37518; AAA49014.1; -.
DR      EMBL; M14278; AAA49035.1; -.
DR      EMBL; M14279; AAA49038.1; -.
DR      EMBL; M14280; AAA49039.1; -.
DR      EMBL; M32732; AAA49011.1; -.
DR      EMBL; M31104; AAA49011.1; JOINED.
DR      EMBL; M32726; AAA49011.1; JOINED.
DR      EMBL; M32727; AAA49011.1; JOINED.
DR      EMBL; M32728; AAA49011.1; JOINED.
DR      EMBL; M32729; AAA49011.1; JOINED.
DR      EMBL; M32730; AAA49011.1; JOINED.
DR      EMBL; M32732; AAA49012.1; -.
DR      EMBL; M31104; AAA49012.1; JOINED.
DR      EMBL; M32726; AAA49012.1; JOINED.
DR      EMBL; M32727; AAA49012.1; JOINED.
DR      EMBL; M32728; AAA49012.1; JOINED.
DR      EMBL; M32729; AAA49012.1; JOINED.
DR      EMBL; M32730; AAA49012.1; JOINED.
DR      EMBL; M31104; AAA49009.1; -.
DR      EMBL; M31104; AAA49010.1; -.
DR      PIR; A35466; A35466.
DR      HSSP; P06401; 1A28.
DR      TRANSFAC; T00698; -.
DR      InterPro; IPR000536; Hormone_rec_lig.
DR      InterPro; IPR000128; Progest_receptor.
DR      InterPro; IPR001723; Stdhrmn_receptor.
DR      InterPro; IPR008946; Str_ncl_receptor.
DR      InterPro; IPR001628; Znf_C4steroid.
DR      Pfam; PF00104; hormone_rec; 1.
DR      Pfam; PF02161; Prog_receptor; 1.
DR      Pfam; PF00105; zf-C4; 1.
DR      PRINTS; PR00398; STRDHORMONER.
DR      PRINTS; PR00047; STROIDFINGER.
DR      ProDom; PD000035; Znf_C4steroid; 1.

```


DR SMART; SM00430; HOLI; 1.
 DR SMART; SM00399; ZnF_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Steroid-binding; Alternative splicing.
 FT DOMAIN 1 420 MODULATING, PRO-RICH.
 FT DNA_BIND 421 486 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 421 441 C4-TYPE.
 FT ZN_FING 457 481 C4-TYPE.
 FT DOMAIN 487 786 STEROID-BINDING.
 FT DOMAIN 48 80 ASP/GLU-RICH (ACIDIC).
 FT VARSPLIC 1 127 Missing (in isoform B and isoform B').
 FT /FTid=VSP_003707.
 FT VARSPLIC 452 458 QHNYLCA -> TISYHCS (in isoform A' and
 FT isoform B').
 FT /FTid=VSP_003708.
 FT VARSPLIC 459 786 Missing (in isoform A' and isoform B').
 FT /FTid=VSP_003709.
 FT CONFLICT 58 58 E -> DD (IN REF. 2).
 FT CONFLICT 480 480 K -> N (IN REF. 2).
 FT CONFLICT 489 489 G -> A (IN REF. 2).
 FT CONFLICT 577 577 R -> T (IN REF. 2).
 FT CONFLICT 642 642 M -> I (IN REF. 2).
 SQ SEQUENCE 786 AA; 85743 MW; 659559950BC45ED9 CRC64;

Query Match 14.9%; Score 272.1; DB 1; Length 786;
 Best Local Similarity 15.9%; Pred. No. 0.2;
 Matches 121; Conservative 54; Mismatches 119; Indels 469; Gaps 28;

Qy 11 SSSTDSP-----RPPPA-----FKYQFVTEPEDEEDEEEDEEEDD 48
 | | : | : || : | : || : | || : || :
 Db 6 SKETRAPSSARDGAVLLQAPPSRGEAEGIDVALDGLLYPRSSDEEEEEEENEEEEEEEP 65
 Qy 49 EDLEELEVLER-----KPAAG-----LSAAAVPPAAAAPLLDFSSD--SVPP 88
 : || | | : | | | | | | | : ||
 Db 66 QQREEEEEEEEDRDCPSYRPGGSLSKDCLDSVLDTF LAPAAHAAPWSLFGPEVPEVPV 125
 Qy 89 AP--RGP----LPAAPPAAPERQPSWERSPAAP-----AP 117
 || || : || | || | || ||
 Db 126 APMSRGPEQKAVDAGPGAPGPSQP----RPGAPLWPGADSLNVAVKARPGPEDASENRAP 181
 Qy 118 SLP-----PAAAVLPSKLPEDDEP----- 136
 || || | : : ||
 Db 182 GLPGAEEERGFFERDAGPGEGGLAPAAAASPAAV---EPGAGQDYLHVPILPLNSAFLAS 237
 Qy 137 ----- 136
 Db 238 RTRQLLDVEAAYDGSFAFGPRSSPSVPAADLAEYGYPPPDGKEGPFAYGEFQSALKIKEEG 297
 Qy 137 ---PARPPP-----PPPAGASPLAE-----PAAP---PST 160
 || || | || | | || ||
 Db 298 VGLPAAPPPFLGAKAAPADFAQPPRAGQEPSLECVLYKAEPPLLPAYGPPAAPDSL PST 357
 Qy 161 PAAP----- 164
 |||
 Db 358 SAAPPGLYSPLGLNGHHQALGFPAAVLKEGLPQLCPPYLG YVRPDTETSQSSQYSFESLP 417

Qy 165 -----KR-----RGS GSSVVDLL-----Y 178
 Db 418 QKICLICGDEASGCHYGVLTGSGCKVFFKRAMEGQHNYLCAGRNDKIRRKNCPCACR 477
 Qy 179 WRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEG 238
 Db 478 LRKCCQAGMVLGGRKFKKLN-----KMKVVRTLDVALQQP--- 512
 Qy 239 HPFRAYLESEVAISEELVQKYS---NSALGHVNSTIKELR----- 275
 Db 513 ----AVLQDE---TQSLTQRLSFSPNQEI PFVPPMISVLRGIEPEVVYAGYDNTK PETPS 565
 Qy 276 -----RLFLVDDLVDLSLKFAVIMWVFTYVGALENGLTTL 308
 Db 566 SLLTSLNHL CERQLLCVVKWSKLLPGFRNLHIDDQITLIQYS---W-----MSL 611
 Qy 309 LILAL-----ISLFSIPVIYERHQVQIDHY 333
 Db 612 MVFAMGWR SYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLSMWQLPQEEFVRLQVSQEEF 671
 Qy 334 L-----GLANKSVKDAM----AKIQAKIPGLKRK 358
 Db 672 LCMKALLLLNTIPLEGLRSQSQFDEMRTSYIRELVKAIGLRQK 714

RESULT 11

CAPU_DROME

ID CAPU_DROME STANDARD; PRT; 1059 AA.
 AC Q24120; Q9VQV8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cappuccino protein.
 GN CAPU OR CG3399.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=96033799; PubMed=7590229;
 RA Emmons S., Phan H., Calley J., Chen W., James B., Manseau L.;
 RT "Cappuccino, a Drosophila maternal effect gene required for polarity
 RT of the egg and embryo, is related to the vertebrate limb deformity
 RT locus.";
 RL Genes Dev. 9:2482-2494(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).

CC -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
 CC -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
 CC -!- SIMILARITY: Belongs to the formin homology family. Cappuccino
 CC subfamily.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U34258; AAC46925.1; -.
 DR EMBL; AE003578; AAF51054.1; -.
 DR PIR; T13286; T13286.
 DR FlyBase; FBgn0000256; capu.
 DR GO; GO:0007304; P:eggshell formation (sensu Insecta); IMP.
 DR GO; GO:0007316; P:pole plasm RNA localization; IMP.
 DR InterPro; IPR000269; CuNH_oxidase.
 DR InterPro; IPR003104; FH2.
 DR InterPro; IPR001265; Formin.
 DR Pfam; PF02181; FH2; 1.
 DR PRINTS; PR00828; FORMIN.


```

Db      870 RTYIAQRRKEGVHPLEIRLPIPEPADVERAAQMDFEVQQQIFDLNKKFLGCKRTTAKVL 929
Qy      312 ALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKI 352
        |   |   |   |   :   ::   ::   |::   |||:   :
Db      930 AA----SRPEIMEPFKSKMEEFVEGADKS----MAKLHQSL 962

```

RESULT 12

MLL4_HUMAN

```

ID      MLL4_HUMAN      STANDARD;      PRT; 2715 AA.
AC      Q9UMN6; O15022; O95836; Q96GP2; Q96IP3; Q9UK25; Q9Y668; Q9Y669;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Myeloid/lymphoid or mixed-lineage leukemia protein 4 (Trithorax
DE      homolog 2).
GN      MLL4 OR TRX2 OR HRX2 OR MLL2 OR KIAA0304.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RA      Angrand P.O., Valvatne H., Jeanmougin F., Adamson A.,
RA      van der Hoeven F., Olsen L., Tekotte H., Huang N., Poch O.,
RA      Lamerdin J., Chambon P., Losson R., Stewart A., Aasland R.;
RT      "Mammalian trithorax- and ASH1-like proteins: putative chromatin
RT      regulators which contain PHD fingers and SET domains.";
RL      Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RA      Lamerdin J.E., McCready P.M., Adamson A.W., Burkhardt-Schultz K.,
RA      Garcia E., Kyle A., Ramirez M., Stilwagen S., Garnes J., Danganan L.,
RA      Bruce R., Quan G., Montgomery M., Ow D., Kobayashi A., Olsen A.O.,
RA      Carrano A.V.;
RT      "Sequence analysis of a 1 Mb region in human 19q13.1.";
RL      Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN      [3]
RP      SEQUENCE OF 111-2715 FROM N.A. (ISOFORM 1).
RC      TISSUE=Leukocyte, and Testis;
RX      MEDLINE=20105772; PubMed=10637508;
RA      Huntsman D.G., Chin S.-F., Muleris M., Batley S.J., Collins V.P.,
RA      Wiedemann L.M., Aparicio S., Caldas C.;
RT      "MLL2, the second human homolog of the Drosophila trithorax gene, maps
RT      to 19q13.1 and is amplified in solid tumor cell lines.";
RL      Oncogene 18:7975-7984(1999).
RN      [4]
RP      SEQUENCE OF 816-2715 FROM N.A. (ISOFORM 1).
RC      TISSUE=Brain;
RX      MEDLINE=97349984; PubMed=9205841;
RA      Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA      Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT      "Prediction of the coding sequences of unidentified human genes. VII.
RT      The complete sequences of 100 new cDNA clones from brain which can
RT      code for large proteins in vitro.";
RL      DNA Res. 4:141-150(1997).

```

RN [5]
 RP SEQUENCE OF 1918-2715 FROM N.A.
 RC TISSUE=Brain, and Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [6]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Bone marrow, and Placenta;
 RX MEDLINE=99339983; PubMed=10409430;
 RA FitzGerald K.T., Diaz M.O.;
 RT "MLL2: A new mammalian member of the trx/MLL family of genes.";
 RL Genomics 59:187-192(1999).
 CC -!- FUNCTION: Possibly acts as a transcriptional regulatory factor.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=Long;
 CC IsoId=Q9UMN6-1; Sequence=Displayed;
 CC Name=2; Synonyms=Truncated;
 CC IsoId=Q9UMN6-2; Sequence=VSP_006668, VSP_006669;
 CC -!- TISSUE SPECIFICITY: Widely expressed. Highest levels in testis.
 CC Also found in brain, bone marrow, heart, muscle, kidney, pancreas,
 CC spleen, thymus, prostate, ovary, intestine, colon, peripheral
 CC blood lymphocytes, and placenta.
 CC -!- DISEASE: Often amplified in pancreatic carcinomas.
 CC -!- SIMILARITY: Belongs to the TRX/MLL family.
 CC -!- SIMILARITY: Contains 1 bromodomain.
 CC -!- SIMILARITY: Contains 1 SET domain.
 CC -!- SIMILARITY: Contains 3 PHD-type zinc fingers.
 CC -!- SIMILARITY: Contains 1 CXXC-type zinc finger.
 CC -!- SIMILARITY: Contains 1 post-SET domain.
 CC -!- CAUTION: This protein was first named MLL2 by Ref.3 and Ref.6.
 CC MLL2 corresponds to another protein located on chromosome 12 (see
 CC AC 014686).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; AJ007041; CAB45385.1; -.
 DR EMBL; AD000671; -; NOT_ANNOTATED_CDS.
 DR EMBL; AB002302; BAA20763.2; -.
 DR EMBL; AF186605; AAD56420.1; -.
 DR EMBL; AF104918; AAD17932.1; -.
 DR EMBL; AF105279; AAD26113.1; -.
 DR EMBL; BC009337; AAH09337.1; -.
 DR EMBL; BC007353; AAH07353.1; -.
 DR EMBL; AF105280; AAD26112.1; -.
 DR MIM; 606834; -.
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003700; F:transcription factor activity; NAS.
 DR GO; GO:0008270; F:zinc ion binding; NAS.
 DR GO; GO:0048096; P:chromatin-mediated maintenance of transcrip. . .; NAS.
 DR InterPro; IPR003889; FYrich_C.
 DR InterPro; IPR003888; FYrich_N.
 DR InterPro; IPR003616; PostSET.
 DR InterPro; IPR001214; SET.
 DR InterPro; IPR002857; Znf_CXXC.
 DR InterPro; IPR001965; Znf_PHD.
 DR Pfam; PF00628; PHD; 3.
 DR Pfam; PF00856; SET; 1.
 DR Pfam; PF02008; zf-CXXC; 1.
 DR SMART; SM00542; FYRC; 1.
 DR SMART; SM00541; FYRN; 1.
 DR SMART; SM00249; PHD; 4.
 DR SMART; SM00508; PostSET; 1.
 DR SMART; SM00317; SET; 1.
 DR PROSITE; PS50868; POST_SET; 1.
 DR PROSITE; PS50280; SET; 1.
 DR PROSITE; PS01359; ZF_PHD_1; 3.
 DR PROSITE; PS50016; ZF_PHD_2; 3.
 KW DNA-binding; Bromodomain; Nuclear protein; Zinc-finger; Metal-binding;
 KW Transcription regulation; Alternative splicing; Repeat.
 FT DNA_BIND 37 44 A.T HOOK (BY SIMILARITY).
 FT DNA_BIND 110 117 A.T HOOK (BY SIMILARITY).
 FT DNA_BIND 357 365 A.T HOOK (BY SIMILARITY).
 FT ZN_FING 959 1005 CXXC-TYPE.
 FT ZN_FING 1201 1252 PHD-TYPE 1.
 FT ZN_FING 1249 1303 PHD-TYPE 2.
 FT ZN_FING 1335 1396 PHD-TYPE 3.
 FT DOMAIN 1449 1471 BROMODOMAIN (DIVERGENT).
 FT DOMAIN 2574 2695 SET.
 FT DOMAIN 2699 2715 POST-SET.
 FT DOMAIN 26 37 POLY-GLY.
 FT DOMAIN 248 255 POLY-PRO.
 FT DOMAIN 362 398 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 402 771 PRO-RICH.
 FT DOMAIN 808 812 POLY-GLN.
 FT DOMAIN 1963 1970 POLY-PRO.
 FT DOMAIN 2251 2259 POLY-PRO.
 FT VARSPLIC 532 582 VSARSSRVIKTPRRFMDEDPPKPPKVEVSPVLRPPITTSP

FT				VPQEPAPVPS -> PLSQSLLLPMTLQLSLSLGQWAAPTTS
FT				ACLDSPWLSPLLLPRCPLTGLQL (in isoform 2).
FT				/FTId=VSP_006668.
FT	VARSPLIC	583	2715	Missing (in isoform 2).
FT				/FTId=VSP_006669.
FT	CONFLICT	834	834	K -> E (IN REF. 6).
FT	CONFLICT	941	941	S -> Y (IN REF. 6).
FT	CONFLICT	1317	1317	E -> Q (IN REF. 6).
FT	CONFLICT	1362	1362	H -> Y (IN REF. 6).
FT	CONFLICT	1438	1438	D -> N (IN REF. 6).
FT	CONFLICT	1918	1920	PLA -> GTR (IN REF. 5; AAH09337).
FT	CONFLICT	2541	2543	DEE -> ARG (IN REF. 5; AAH07353).
FT	CONFLICT	2622	2622	D -> H (IN REF. 6).
SQ	SEQUENCE	2715	AA; 293511	MW; C0615B981BBEB7BF CRC64;

Query Match 14.1%; Score 257.6; DB 1; Length 2715;
 Best Local Similarity 12.6%; Pred. No. 12;
 Matches 114; Conservative 30; Mismatches 100; Indels 664; Gaps 22;

Qy	2	EDIDQSSLVSS-----STDSP-----	17
		: : :	
Db	130	EDVAPSSLRSALRSQRGRAPRGRGRKHKTPLPPPLRLADVAPTPPKTPARKRGEETGTERM	189
Qy	18	-----	17
Db	190	VQALTELLRRAQAPQAPRSRACEPSTPRRSRGRPPGRPAGPCRRKQQAQVVAEAAVTIPK	249
Qy	18	PRPPP-----AFKYQFVT	30
		: :	
Db	250	PEPPPPVVPVKHQTGSWKCKEGPGPGPTPRRGQSSRGGRGRGRGRGGGLPFVIKQFVS	309
Qy	31	-----EPEDEEDEEEEEDE	44
		: : : :	
Db	310	RAKKVKMGQLSLGLESGQGQGHESWQDVPQRRVSGQGGSPCWKKQEQLDDEEEEEKK	369
Qy	45	EEDDEDLEEEVLER-----KPAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAP	98
		::: :	
Db	370	EEEEKDKEGEEKEERAVAEEMMPAAEKEEAKLPP---PPL-----TPPAPSPPPPLPP	419
Qy	99	PAA-----	101
		:	
Db	420	PSTSPPPPLCPPPPPPVSPPPPLSPPPPPAQQEEQEESSPPVVPATCSRKRGRPPLTPSQR	479
Qy	102	-----	101
Db	480	AEREAARAGPEGTSPPTPTPSTATGGPPEDSPTVAPKSTTFLKNIRQFIMPVVSARSSRV	539
Qy	102	-----PERQPSWERSPA-----APAPSLPPAAAVLPS---KLPE	132
		:	
Db	540	IKTPRRFMDEDPPKPPKVEVSPVLRPPITTSPPVPQEPAPVPS-PPRAPTPPSTPVPLPE	598
Qy	133	-----DDEPPARPPP-----	142
Db	599	KRRSILREPTFRWTSLTRELPPPPPPAPPPPPAPSPPPAPATSSRRPLLLRAPQFTPSEAH	658
Qy	143	-----PPPAGASPLAEPAAAPS--TPAAPKRRGSGSSVVDLLYWRDIKKTGVVFGA	191
		: : : :	


```

Db      659 LKIYESVLTPPPLGAPEAPEPEPPPADDSPAEPPEPRAVGRT----- 699
Qy      192 SLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPPFRAYLESEVAI 251
      ||| | : | | | | : | | |
Db      700 ---NHLSSLPRFAPVVTTPVKA-----EVSPHGAP----- 725
Qy      252 SEELVQKYSNSALGHVNSTIKELRRLEFLVDDLVDSLKFAVLMWVFTYVGALENG----- 305
      || ||
Db      726 -----ALSNGPQTQAQ 736
Qy      306 -----LTLL 309
      : ||
Db      737 LLQPLQALQTQLLPQALPPPQPQLQPPPSFQMPPLEKARIAGVGSPLPSGVEEKMFSL 796
Qy      310 ILALISLFSI-----PVIYERHQVQ----- 329
      | : || | || | | :
Db      797 KRAKVQLFKIDQQQQQKVAASMPSPGGQMEEVAGAVKQISDRGPVRSEDESVEAKRERP 856
Qy      330 -----IDHY-----LGLANKSVKDAMAKIQA----- 350
      | | || | | : : : |
Db      857 SGPESPVQGPRIKHVCRHAAVALGQARAMVPEDVPRLSALPLRDRQDLATEDTSSASETE 916
Qy      351 KIPGLKRK 358
      : | | :
Db      917 SVPSRSRR 924

```

RESULT 13

FM14_MOUSE

```

ID FM14_MOUSE STANDARD; PRT; 1206 AA.
AC Q05859;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Formin 1 isoform IV (Limb deformity protein).
GN FMN OR LD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=92112033; PubMed=1339380;
RA Grusby-Jackson L., Kuo A., Leder P.;
RT "A variant limb deformity transcript expressed in the embryonic mouse
RT limb defines a novel formin.";
RL Genes Dev. 6:29-37(1992).
CC -!- FUNCTION: Is important in the morphogenesis of limb and may have a
CC function in differentiated cells or be involved in maintaining
CC specific differentiated states.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Comment=Additional isoforms seem to exist;
CC Name=IV;
CC IsoId=Q05859-1; Sequence=Displayed;
CC Name=IA;

```

```

CC      IsoId=Q05860-1; Sequence=External;
CC      Name=IB;
CC      IsoId=Q05860-2; Sequence=External;
CC      Name=II;
CC      IsoId=Q05860-3; Sequence=External;
CC      Name=III;
CC      IsoId=Q05860-4; Sequence=External;
CC      -!- TISSUE SPECIFICITY: It is found throughout the embryo but has a
CC          functional role only in the kidney and limb.
CC      -!- DEVELOPMENTAL STAGE: This is the isoform found in the apical
CC          ectodermal ridge and the mesenchymal compartment of the developing
CC          limb bud.
CC      -!- PTM: PHOSPHORYLATED ON SERINE AND POSSIBLY THREONINE RESIDUES.
CC      -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
CC      -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
CC      -!- SIMILARITY: Belongs to the formin homology family. Cappuccino
CC          subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X62379; CAA44244.1; -.
DR      PIR; S24407; S24407.
DR      MGD; MGI:101815; Fmn.
DR      InterPro; IPR003104; FH2.
DR      InterPro; IPR001265; Formin.
DR      Pfam; PF02181; FH2; 1.
DR      PRINTS; PR00828; FORMIN.
KW      Nuclear protein; Developmental protein; Alternative splicing;
KW      Phosphorylation; Coiled coil.
FT      DOMAIN      418      443      COILED COIL (POTENTIAL).
FT      DOMAIN      497      566      COILED COIL (POTENTIAL).
FT      DOMAIN      644      744      FH1 (PRO-RICH).
FT      DOMAIN      759      1164     FH2.
FT      DOMAIN      1043     1116     COILED COIL (POTENTIAL).
FT      DOMAIN      635      638      POLY-SER.
FT      DOMAIN      751      755      POLY-SER.
SQ      SEQUENCE    1206 AA;  133464 MW;  4DFB38CB52BD8EE7 CRC64;

```

```

Query Match          14.1%;  Score 257.5;  DB 1;  Length 1206;
Best Local Similarity 14.1%;  Pred. No. 1.8;
Matches    95;  Conservative    39;  Mismatches    95;  Indels    445;  Gaps    22;

```

```

Qy      13 STDSPPPRPPP-----AFKYQFV----- 29
          |||      |          |||  |:
Db      373 STDQESHKSPRDAHVQGGQVKARTPETALEAFKALFIRPPKKGSTADTSELEALKRKMKH 432

Qy      30 -----TEPEDEEDEEEEEDE---EEDD----- 48
          : | |      : :  |:  |:||
Db      433 EKESLRAVFERSKSRPADSPSPDKSPDQSPTEQDDRTPGRLQAVWPPPKTKDTEEKVGLK 492

Qy      49 -----EDLEELEVLER----- 59

```

```

      |  ||:| |:
Db      493 YTEAEYQAAILHLKREHKKEIETLQAQFELKTFHIRGEHALVTARLEEAIENLKQOLEKR 552
Qy      60 -----KPAAGLSAAA----- 69
      ||  ||
Db      553 REGCEEMRDVCISTDDDCSPKAFRNVCIQTDRETFLKPCDAESKATRSSHQIVPKKLTISL 612
Qy      70 -----VPPAAAAPLLDFSSDSVPPAP--- 90
      ||  ||  :|||
Db      613 TQLSPSKDSKDIHAPFQTREGTSSSSQOKISPPAPPTPPPLPPPL-----IPPPPLP 665
Qy      91 --RGPLPAAPPAAPERQPSWERSPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPP--- 145
      |||| ||| |  | :| | ||  | : :| | || |||||
Db      666 PGLGPLPPAPPPIPPV-----CPVSPPPPPPPPPP---PTPVPPSDGPPPPPPPPPLPN 715
Qy      146 -----AGASP-----LAEPAAP-----PSTPAAPKRRG 168
      :| |  || | |  || | |
Db      716 VLALPNSGGPPPPPPPPPPPPGLAPPPPPGLSFGLSSSSSQYPRKPAIEPSCPMKP---- 771
Qy      169 SGSSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 228
      |||
Db      772 -----LYWTR----- 776
Qy      229 IQAIQKSDEGHP-----FRAYLESEVAISEELVQ----- 257
      ||  || : |  | | | | : : |
Db      777 IQINDKSQDAAPTLLWDSLEEPHIRTSEFEYLFSDKDTTQQKKKPLSEAYEKKNKVKKIIK 836
Qy      258 ----KYSNSA---LGHVNSTIKELRR-LFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLL 309
      | | :  : : : :| : : : | |||
Db      837 LLDGKRSQTVGILISLHLEMKDIQQAIFTVDD----- 869
Qy      310 ILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQ----- 349
      : : | : : : ||  | : : | : || :
Db      870 --SVVDLETLAALYE-----NRAQEDELTKIRKYYETSKEEDLKLLDKPEQF 914
Qy      350 ----AKIPGLKRKA 359
      | : ||  : |
Db      915 LHELAQIPNFAERA 928

```

RESULT 14

MY15_HUMAN

ID MY15_HUMAN STANDARD; PRT; 3530 AA.

AC Q9UKN7;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Myosin XV (Unconventional myosin-15).

GN MYO15A OR MYO15.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.

RX MEDLINE=20021762; PubMed=10552926;

RA Liang Y., Wang A., Belyantseva I.A., Anderson D.W., Probst F.J.,

RA Barber T.D., Miller W., Touchman J.W., Jin L., Sullivan S.L.,
 RA Sellers J.R., Camper S.A., Lloyd R.V., Kachar B., Friedman T.B.,
 RA Fridell R.A.;
 RT "Characterization of the human and mouse unconventional myosin XV
 RT genes responsible for hereditary deafness DFNB3 and shaker 2.";
 RL Genomics 61:243-258(1999).
 RN [2]
 RP PARTIAL SEQUENCE FROM N.A., AND VARIANTS DFNB3 TYR-2111 AND PHE-2113.
 RX MEDLINE=98267311; PubMed=9603736;
 RA Wang A., Liang Y., Fridell R.A., Probst F.J., Wilcox E.R.,
 RA Touchman J.W., Morton C.C., Morell R.J., Noben-Trauth K., Camper S.A.,
 RA Friedman T.B.;
 RT "Association of unconventional myosin MYO15 mutations with human
 RT nonsyndromic deafness DFNB3.";
 RL Science 280:1447-1451(1998).
 CC -!- FUNCTION: Myosins are actin-based motor molecules with ATPase
 CC activity. Unconventional myosins serve in intracellular movements.
 CC Their highly divergent tails are presumed to bind to membranous
 CC compartments, which would be moved relative to actin filaments (By
 CC similarity). May play a role in the formation or maintenance of
 CC the actin-rich structures of the inner ear sensory hair cells.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- TISSUE SPECIFICITY: Highly expressed in pituitary. Also expressed
 CC at lower levels in adult brain, kidney, liver, lung, pancreas,
 CC placenta and skeletal muscle. Not expressed in brain. In the
 CC pituitary, highly expressed in anterior gland cells.
 CC -!- DISEASE: Defects in MYO15A are the cause of autosomal recessive
 CC nonsyndromic deafness type 3 (DFNB3) [MIM:600316].
 CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
 CC -!- SIMILARITY: Contains 1 FERM domain.
 CC -!- SIMILARITY: Contains 3 IQ domains.
 CC -!- SIMILARITY: Contains 1 MyTH4 domain.
 CC -!- SIMILARITY: Contains 1 SH3 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF144094; AAF05903.1; -.
 DR EMBL; AF051976; -; NOT_ANNOTATED_CDS.
 DR PIR; A59266; A59266.
 DR HSSP; P10587; 1BR2.
 DR Genew; HGNC:7594; MYO15A.
 DR MIM; 602666; -.
 DR MIM; 600316; -.
 DR GO; GO:0007605; P:hearing; TAS.
 DR InterPro; IPR000299; Band_4.1.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR001609; myosin_head.
 DR InterPro; IPR000857; MyTH4.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00612; IQ; 3.
 DR Pfam; PF00063; myosin_head; 1.

DR Pfam; PF00784; MyTH4; 2.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00295; B41; 1.
 DR SMART; SM00015; IQ; 3.
 DR SMART; SM00242; MYSc; 1.
 DR SMART; SM00139; MyTH4; 2.
 DR SMART; SM00326; SH3; 1.
 DR PROSITE; PS00660; FERM_1; FALSE_NEG.
 DR PROSITE; PS00661; FERM_2; FALSE_NEG.
 DR PROSITE; PS50057; FERM_3; 1.
 DR PROSITE; PS50096; IQ; 3.
 DR PROSITE; PS50002; SH3; FALSE_NEG.
 KW Myosin; ATP-binding; Actin-binding; Coiled coil; Repeat; SH3 domain;
 KW Calmodulin-binding; Disease mutation; Deafness.
 FT DOMAIN 1 1887 MYOSIN HEAD-LIKE.
 FT DOMAIN 1888 2029 NECK OR REGULATORY DOMAIN.
 FT DOMAIN 2030 3530 TAIL.
 FT DOMAIN 1323 1350 COILED COIL (POTENTIAL).
 FT DOMAIN 1792 1799 ACTIN-BINDING (POTENTIAL).
 FT DOMAIN 1902 1924 IQ 1.
 FT DOMAIN 1925 1954 IQ 2.
 FT DOMAIN 1955 1976 IQ 3.
 FT DOMAIN 2867 2953 SH3.
 FT DOMAIN 3209 3530 FERM.
 FT NP_BIND 1315 1322 ATP (POTENTIAL).
 FT VARIANT 2111 2111 N -> Y (in DFNB3; family from Bengkulu).
 FT /FTId=VAR_010303.
 FT VARIANT 2113 2113 I -> F (in DFNB3; Indian family).
 FT /FTId=VAR_010304.
 SQ SEQUENCE 3530 AA; 395171 MW; 3D103923D4BCBE4A CRC64;

Query Match 14.1%; Score 256.4; DB 1; Length 3530;
 Best Local Similarity 10.1%; Pred. No. 26;
 Matches 145; Conservative 54; Mismatches 124; Indels 1106; Gaps 37;

QY 17 PPRPP-----PAFKYQFV-----TEPEDEEDEEE----- 40
 || | || | : | || | | :
 Db 394 PPEVPYFYPEESASAFVYPWPVPPPIPSPHNPHYAHMDDIAELEEPEDAGVERQGTSFRLP 453
 QY 41 -----EEDEE 45
 : |||
 Db 454 SAAFFEQQGMDKPARSKLSLIRKFRLFP RPQVKLFGKEKLEVLPPLPSLDIPLPLGDADEE 513
 QY 46 EDDLEEELEVL----- 57
 ||:|:| : :
 Db 514 EDEELPPVSAVPYGHFWGFLTPRQRNLQRALSAFGAHRGLGFGPEFGRVPRPATSLA 573
 QY 58 -----ERKPAAGL----- 65
 |:| | |
 Db 574 RFLKKTLSKKPIARLRGSQKTRAGGPAVREAAAYKRFGYKLAGMDPEKPGTPIVLRRAQP 633
 QY 66 ----- 65
 Db 634 RARSSNDARRPPAPQPAPRTLSHWSALLSPVPVPRPPSSGPPAPPLSPALSGLPRPAS 693
 QY 66 -----SAAAVPPAAAA----- 76

```

      : | | | | |
Db      694 YGSLRRHPPPPWAAPAHVPPAPQASGWAFVEPPAVSPEVPPDLLAFPGPRPSFRGSRRRGA 753
Qy      77 -----PLLDfSSDSVPPAPR-----GPL--PAAPP 99
      | | : | | | : | | | |
Db      754 AFGFPGASPRASRRRAWSPLASPQPSLRSSPGLGYCSPLAPPSPQLSLRTGPFQPPFLPP 813
Qy      100 A-----APERQ-----PSWERSPAAPAPSLPP-----AAAVLPSKL----- 130
      | | | | | | : | | | | : | | | : |
Db      814 ARRPRSLQESPAPRRAAGRLGPPGSPLPGSPRPPSPPLGLCHSPRRSSNLNLP SRLPHTWR 873
Qy      131 -----PEDDE----- 135
      | | | |
Db      874 RLSEPPTRAVKPQVRLPFHRPPRAGAWRAPLEHRESPREPEDSETPWTVPPLAPSWDMD 933
Qy      136 -PPARPPPPPPAGA-----SPLAEPAAP----- 157
      | | | | | | | | | | : | |
Db      934 PPTQRPPSPWPGGAGSRRGFSRPPVPENPFLQLLGPVPSPTLQPEDPAADMTRVFLGRH 993
Qy      158 -----PSTPAAPK----- 165
      | | | | |
Db      994 HEPGPGQLTKSAGPTPEKPEEEATLGDPQLPAETKPTPAPPKDVTPPKDITPPKDVLP 1053
Qy      166 -----RR----- 167
      | |
Db      1054 QKTLRPSLSYPLAACDQTRATWPPWHRWGTL PQAAAPLAPIRAPEPLPKGGERRQAAPGR 1113
Qy      168 ----- 167
Db      1114 FAVVMPRVQKLSSFQRVGPATLKPQVQPIQDPKPRACSLRWSC LWRADAYGPWPRVH 1173
Qy      168 -----GSGSS-----VVDLLYWRDI 182
      | | : | | : | :
Db      1174 PQSCHLGPGAACLSLRGSWEEVGPPSWRNKMHSIRNLP SMRFREQHGEGDGVEDMTQLEDL 1233
Qy      183 KKTGVVFGASLFLLLSLTVFSIVSVT-----AYIALALLSVT--ISFRIY----- 225
      : : | | | : : | | | : | | | |
Db      1234 QET-----TVLSNLKIRFERNLIYTYIGSILVSVNPYQMFGIYGPEQVQY 1279
Qy      226 -----K 226
      :
Db      1280 NGRALGENPPHLFAVANLAFAKMLDAKQNCIIISGESGSGKTEATKLILRYLAAMNQKR 1339
Qy      227 GVIQAIQ-----KSDEGHPPFRAYLE-----SEVAISEELVQK--- 258
      | : | | : : | | : : | | : : |
Db      1340 EVMQQIKILEATPLLESFGNAKTVRNDNSSRFGKFVEIFLEGGVISGAITSQYLLEKSRI 1399
Qy      259 -----YSNSA----- 263
      | |
Db      1400 VFQAKNERNYHIFYELLAGLPAQLRQAFSLQEAETYYYNQGGNCEIAGKSDADDFRLL 1459
Qy      264 -----LGH----- 266
      | | :
Db      1460 AAMEVLGFSSDQDSIFRILASILHLGNVYFEKYETDAQEVASVVSAREIQAVAELLQIS 1519
Qy      267 -----VNSTIKELRRLF-----LVD--DLVDSLKFAVLM-WVFTYVGAL--- 302
      | | : : | | | : : : : | | | |

```

```

Db      1520 PEG LQKAITFKVTETMRE--KIFTPLTVESAVDARDAlAKVLYALLFSWLITRVNALVSP 1577
Qy      303 -----FNGLTLLIL-----ALISLFSIPVI-----YERHQVQ--- 329
          || | : | ||: | | | :
Db      1578 RQDTLSIAILDYGFEDLSFNSFEQLCINYANENLQYLFNKIVFQEEQEEYIREQIDWQE 1637
Qy      330 ----- 329
Db      1638 ITFADNQPRINLISLKPYGILRILDDQCCFPQATDHTFLQKCHYHHGANPLYSKPKMPLP 1697
Qy      330 ---IDHYLG LA----- 337
          | | | |
Db      1698 EFTIKHYAGKVITYQVHKFLDKNHDQVRQDVLDFVRSRTRVVAHLFSSHAPQAAPQRLGK 1757
Qy      338 -----NKS VKDAMAKIQ-----AKIPGL 355
          : | : | : | | |
Db      1758 SSSVTRLYKAHTVAAKFQQLDLVEKMERCNPLFMRCLKPNHKKEPGL 1806

```

RESULT 15

FMN1_MOUSE

```

ID      FMN1_MOUSE      STANDARD;      PRT; 1468 AA.
AC      Q05860;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Formin 1 isoforms I/II/III (Limb deformity protein).
GN      FMN OR LD.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney, and Testis;
RX      MEDLINE=90363291; PubMed=2392150;
RA      Woychik R.P., Maas R.L., Zeller R., Vogt T.F., Leder P.;
RT      "'Formins': proteins deduced from the alternative transcripts of the
RT      limb deformity gene.";
RL      Nature 346:850-853(1990).
RN      [2]
RP      ALTERNATIVE SPLICING.
RX      MEDLINE=97224459; PubMed=9119367;
RA      Wang C.C., Chan D.C., Leder P.;
RT      "The mouse formin (Fmn) gene: genomic structure, novel exons, and
RT      genetic mapping.";
RL      Genomics 39:303-311(1997).
RN      [3]
RP      PHOSPHORYLATION.
RX      MEDLINE=93296176; PubMed=8516300;
RA      Vogt T.F., Jackson-Grusby L., Rush J., Leder P.;
RT      "Formins: phosphoprotein isoforms encoded by the mouse limb deformity
RT      locus.";
RL      Proc. Natl. Acad. Sci. U.S.A. 90:5554-5558(1993).
CC      -!- FUNCTION: Is important for the morphogenesis of limb and kidney
CC      and may have a function in differentiated cells or may be involved
CC      in maintaining specific differentiated states.

```

CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=5;
 CC Comment=Additional isoforms seem to exist;
 CC Name=IA;
 CC IsoId=Q05860-1; Sequence=Displayed;
 CC Name=IB;
 CC IsoId=Q05860-2; Sequence=VSP_001569;
 CC Name=II;
 CC IsoId=Q05860-3; Sequence=VSP_001570;
 CC Name=III;
 CC IsoId=Q05860-4; Sequence=VSP_001571, VSP_001572;
 CC Name=IV;
 CC IsoId=Q05859-1; Sequence=External;
 CC -!- TISSUE SPECIFICITY: It is present in the adult kidney, testis,
 CC limb, ovary, brain, small intestine, salivary gland and harderian
 CC gland. It is present throughout the embryo.
 CC -!- DEVELOPMENTAL STAGE: In the developing limb bud, the protein is
 CC expressed in the apical ectodermal ridge and the mesenchymal
 CC compartment, predominantly in the posterior region. During kidney
 CC morphogenesis, expression is initially restricted to the
 CC epithelial compartment of the pronephros and mesonephros.
 CC -!- PTM: PHOSPHORYLATED ON SERINE AND POSSIBLY THREONINE RESIDUES.
 CC -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
 CC -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
 CC -!- SIMILARITY: Belongs to the formin homology family. Cappuccino
 CC subfamily.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X53599; CAA37668.1; -.
 DR PIR; S11515; S11515.
 DR MGD; MGI:101815; Fmn.
 DR InterPro; IPR003104; FH2.
 DR InterPro; IPR001265; Formin.
 DR Pfam; PF02181; FH2; 1.
 DR PRINTS; PR00828; FORMIN.
 DR SMART; SM00498; FH2; 1.
 KW Nuclear protein; Developmental protein; Alternative splicing;
 KW Phosphorylation; Coiled coil.
 FT DOMAIN 723 792 COILED COIL (POTENTIAL).
 FT DOMAIN 870 970 FH1 (PRO-RICH).
 FT DOMAIN 985 1426 FH2.
 FT DOMAIN 1305 1378 COILED COIL (POTENTIAL).
 FT DOMAIN 198 203 POLY-SER.
 FT DOMAIN 861 864 POLY-SER.
 FT DOMAIN 885 892 POLY-PRO.
 FT DOMAIN 911 925 POLY-PRO.
 FT DOMAIN 929 940 POLY-PRO.
 FT DOMAIN 951 962 POLY-PRO.
 FT DOMAIN 966 970 POLY-PRO.

FT	DOMAIN	977	981	POLY-SER.
FT	VARSPPLIC	1252	1287	Missing (in isoform IB).
FT				/FTId=VSP_001569.
FT	VARSPPLIC	625	722	Missing (in isoform II).
FT				/FTId=VSP_001570.
FT	VARSPPLIC	626	627	IA -> SV (in isoform III).
FT				/FTId=VSP_001571.
FT	VARSPPLIC	628	1468	Missing (in isoform III).
FT				/FTId=VSP_001572.
SQ	SEQUENCE	1468 AA;	163809 MW;	EF2FB1E9CA9DAF43 CRC64;

Query Match 14.0%; Score 254.9; DB 1; Length 1468;
 Best Local Similarity 14.5%; Pred. No. 3.5;
 Matches 97; Conservative 40; Mismatches 103; Indels 431; Gaps 23;

Qy	2	EDIDQSSLVSSSTDSPPRP--PPA-----FKYQ-----	27
		: : : :	
Db	602	EEASEKGLGPEKITAPPQHQLPPGIASEGFPDNEKEQTAKDLPNKDGWVWPGYRAGPP	661
Qy	28	--FVTEPEDEEDEEEEE-----EDEEEDD-----	48
		: : : :	
Db	662	CPFLLHEEKEKTSRSELYLDLNPDSPTQDDRTPGRLQAVWPPPKTKDTEEKVGLKYTE	721
Qy	49	-----EDLEELEVLER-----	59
		: :	
Db	722	AEYQAAIHLHKREHKEEIEITLQAQFELKTFHIRGEHALVTARLEEAIENLKQOLEKRREG	781
Qy	60	-----KPAAGLSAAA-----	69
Db	782	CEEMRDVCISTDDDCSPKAFRNVCIQTDRETFLKPCDAESKATRSSQIVPKKLTISLTQL	841
Qy	70	-----VPPAAAAPLLDFSSDSVPPAP-----R	91
		:	
Db	842	SPSKDSKDIHAPFQTREGTSSSSSQKISPPAPPTPPPLPPPL-----IPPPPLPPGL	894
Qy	92	GPLPAAPPAAPERQPSWERSPAAPAPSLPPAAVLPSKLPEDDEPPARPPPPPP-----	145
		: : :	
Db	895	GPLPPAPPPIPPV-----CPVSPPPPPPPPPP---PTPVPPSDGPPPPPPPPPLPNVLA	944
Qy	146	---AGASP-----LAEPAAP-----PSTPAAPKRRGSGS	171
		:	
Db	945	LPNSGGPPPPPPPPPPPPGLAPPPPPGLSFGLSSSSSQYPRKPAIEPSCPMKP-----	997
Qy	172	SVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIIQA	231
Db	998	-----LYWTR-----IQI	1005
Qy	232	IQKSDEGHP-----FRAYLESEVAISEELVQ-----	257
		: :	
Db	1006	NDKSQDAAPTLDWSLEEPHIRTSEFEYLFSDTTQKKKPLSEAYEKKNKVKKIIKLLD	1065
Qy	258	-KYSNSA---LGHVNSTIKELRR-LFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILA	312
		: : : : : :	
Db	1066	GKRSQTVGILISSLHLEMKDIQQAIFTVDD-----S	1096
Qy	313	LISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQ-----	349
		:: :: : :: : : :	

Db 1097 VVDLETLAALYE-----NRAQEDELTKIRKYYETSKEEDLKLLDKPEQFLHE 1143
Qy 350 -AKIPGLKRKA 359
|:| | :|
Db 1144 LAQIPNFAERA 1154

Search completed: September 29, 2004, 18:52:07
Job time : 29.9792 secs